Characteristics and chemical properties of *Escherichia coli* Curli

Von der Fakultät für Lebenswissenschaften

der Technischen Universität Carolo-Wilhelmina

zu Braunschweig

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

genehmigte

D i s s e r t a t i o n

von Puwei Yuan
aus Shannxi / China

1. Referentin: Dr. Christiane Ritter
2. Referent: apl. Prof. Dr. Michael Hust
3. Referentin: Prof. Dr. Susanne Engelmann
eingereicht am: 09.02.2015
mündliche Prüfung (Disputation) am: 08.06.2015

Druckjahr 2015

Vorveröffentlichungen der Dissertation

Teilergebnisse aus dieser Arbeit wurden mit Genehmigung der Fakultät für Lebenswissenschaften, vertreten durch die Mentorin der Arbeit, in folgenden Beiträgen vorab veröffentlicht:

Tagungsbeiträge:

Puwei Yuan, Tobias Schubeis, Mumdooh Ahmed and Christiane Ritter. Solid-state NMR of Curli amyloid fibrils from *Escherichia coli* (Poster). 4th Bio-NMR Annual User Meeting, Warsaw Poland (2014)

Puwei Yuan and Christiane Ritter. Characteristics and chemical properties of amyloid fibrils Curli (poster).6th international PhD Symposium, Braunschweig (2013)

Puwei Yuan and Christiane Ritter. Characteristics and chemical properties of amyloid fibrils Curli (poster).5th international PhD Symposium, Braunschweig (2012)

Acknowledgements

Three and half years of my doctoral is nearing completion. I hereby sincerely thank all the people who helped me during my doctoral.

First of all, I would like to thank my doctoral advisor Dr. Christiane Ritter. She helped me to imperceptibly adapt the german way of scientific research. She's erudition both in amyloid and NMR fields gave me a huge help to study on amyloid Curli. I am grateful that she provided very good experimental conditions and creative environment in her group to help me to complete my doctoral research. Without her meticulous and patient instruction, I would not finish my thesis smoothly.

I would also like to thank Dr. Thorsten Luehrs for kindly providing Selective Shearing Amplification array of his genius invention that helped me to reveal the secreat of mixed CsgA-CsgB fibrils.

I would also like to thank Prof. Dr. Manfred Rohde, a very friendly and rich experience scientist. He gave me a lot of advices and helped me to take many pretty cool fibrils TEM images. Some of nice TEM images were presented in my thesis.

I am also grateful for my thesis committee members of Prof. Dr. Wulf Blankenfeldt, Dr. Joop van den Heuvel and Prof. Dr. Melanie M. Brinkmann for providing scientific discussions and constructive advices of my project.

It is my great pleasure to thank Dr. Lichun He for his friendly help with teaching on NMR data processing & analysis and recording solution NMR spectra; He and his wife Ms. Lisha Zha also gave me many productive suggestions and experimental assistance to help me to finish my project.

I also sincerely thank Dr. Gang Pei who helped me to obtain the high-quality western-bloting results.

I would also like to thank Dr. Gang Zhao who gave me many help on data analysis using Metlab.

My sincere thanks also go to all the lab members: Dr. Madhu Nagaraj, Dr. Mumdooh Ahmed, Dr. Johannes Spehr and Dr. Tobias Schubeis. Thanks for their discussions, suggestions and friendly environment in the lab.

I would also like to thank all my friends in HZI and Braunschweig, for helping in my daily life and making me not feel alone in Germay during three and half years.

Last but not least, my sincere thanks to my parents and relatives for their love and huge encouragement.
Again, I sincerely thank all of you.

Puwei Yuan

28 10 2014
In Braunschweig, Germany

Contents

Abbreviations

aa	Amino acid
Å	Ångström
CD	Circular dichroism
CR	Congo red
CsgX	Curli specific gene X
kDa	Kilo Dalton
E. coli	*Escherichia coli*
EM	Electron microscopy
DMSO	Dimethyl sulfoxide
EDTA	Ethylenediaminetetraacetic acid
FA	Formic acid
H/D	Hydrogen/Deuterium
HSQC	Heteronuclear single quantum coherence
FT-IR	Fourier transforms infrared HRP horseradish peroxidase
IPTG	Isopropyl-β-D-thiolgalactoside
KPi	KH_2PO_4/K_2HPO_4
NMR	Nuclear magnetic resonance
OD600	Optical density at 600 nm
PAGE	Polyacrylamid gel electrophoresis
PCR	Polymerase chain reaction
SDS	Sodium dodecyl sulphate
Sec	Secretory TAE
TBS	Tris buffered saline
TEM	Transmission electron microscopy
TEMED	N, N, N', N' – Tetramethylendiamin
TFA	Trifluoroacetic acid
ThT	Thioflavin T
Tris	Tris (hydroxymethyl) aminomethane
UTI	Urinary tract infection

Zusammenfasung

Curli ist der Hauptproteinbestandteil einer komplexen, extrazellulären Matrix, die von vielen Enterobacteriaceae gebildet wird. Es ist an der Oberflächenadhäsion, an Zellaggregation und an der Biofilmbildung beteiligt. Darüber hinaus ist es ein effizienter Auslöser von Entzündungsreaktionen im Wirt. Bei Curli handelt es sich um Amyloidfibrillen mit einem Durchmesser von sechs bis zwölf Nanometern und einer Länge von mehreren Mikrometern. Die Hauptuntereinheit von Curli ist CsgA, während die Nebenkomponente CsgB als Nukleator fungiert. Dabei ist der N-Terminus von CsgB für die Nukleation von CsgA verantwortlich, während sein C-Terminus eine Rolle bei der Selbstfaltung von CsgB spielt.

Hochaufgelöste Strukturdaten zu Curli-Fibrillen stellen eine Voraussetzung für das Verständnis von Aufbau und Funktion dieser Einheiten dar. Außerdem könnten solche Daten Einblicke in die Unterschiede zwischen krankheitsassoziierten und funktionalen Amyloiden geben. In der vorliegenden Arbeit wurde der Unterschied struktureller Eigenschaften von nativ produziertem Curli und *in vitro* hergestellten Fibrillen von CsgA und/oder CsgB untersucht. Das bakterielle Amyloid Curli bietet hiermit die seltene Gelegenheit, die strukturellen Eigenschaften von rekombinant und natürlich hergestellten Amyloidfibrillen zu untersuchen. Dazu wurde ein Protokoll entwickelt, um ausreichende Mengen an ^{15}N,^{13}C-markiertem, nativen Curli aus *Escherichia coli* MC4100 für die Analyse mittels Festkörper-NMR-Spektroskopie zu exprimieren und zu reinigen. Ein 2D ^{13}C-^{13}C-Festkörper-NMR-Spektrum von nativem Curli zeigte starke Ähnlichkeit zu solchen Spektren, die von rekombinanten CsgA-Fibrillen aufgenommen wurden. Dieser Befund impliziert, dass die Tertiärstruktur von CsgA-Fibrillen große Ähnlichkeit mit der von nativem Curli besitzt.

CsgB kann als Nukleator für CsgA dienen, um gemischte CsgA-CsgB-Fibrillen mit beliebigem Massenverhältnis zu erzeugen. Solche gemischten Fibrillen hatten hier eine hochgeordnete, amyloidähnliche Struktur, die nativem Culi ähnlich ist. *In vitro* hergestellte, gemischte Fibrillen wurden mittels Immungold-TEM, kontrollierter Fragmentierung durch Scherkraft in Kombination mit Nukleationsversuchen und H/D-Austausch-NMR-Spektroskopie untersucht. Immungold-TEM-Aufnahmen konnten dabei eindeutig belegen, dass CsgB nicht nur an den Enden, sondern auch zwischen den einzelnen CsgA-Molekülen der gemischten Fibrillen vorkommt. Sowohl die N-terminalen als auch die C-terminalen Regionen von CsgB interagieren dabei mit benachbarten CsgA-Moleülen. Die Fragmentierung von gemischten CsgA-CsgB-Fibrillen erforderte im Vergleich mit reinen

CsgA-Fibrillen anscheinend leicht geringere Scherkräfte. Dies könnte darauf zurückzuführen sein, dass die eingestreuten CsgB-Einheiten die Gesamtstabilität der Fibrillen verringern. H/D-Austauschdaten haben gezeigt, dass eingestreuten CsgB bildeten eine geordnetere β-Faltblattstruktur im gemischten Fibrillen.

Summary

Curli are the major proteinaceous component of a complex extracellular matrix produced by many enterobacteriaceae. They are involved in adhesion to surfaces, cell aggregation, and biofilm formation. They are also potent inducers of the host inflammatory response. Curli are amyloid fibers 6-12 nm in diameter and several μm in length. CsgA is the major subunit of Curli, while CsgB is the minor nucleator subunit. The N-terminus of CsgB is involved in the nucleation of CsgA while the C-terminus is important for self-folding of CsgB.

High-resolution structural information of Curli fibers will be a prerequisite to gain insights into its assembly and functions, and might help to understand the differences between disease-related and functional amyloids. In this work, I aimed to compare to the structural properties of natively produces Curli with those of *in-vitro* generated fibrils composed of CsgA and/or CsgB. The bacterial amyloid Curli offers therefore the unique opportunity to compare the structural properties of recombinant and natively produce amyloid fibrils. I developed the protocol to express and purify sufficient amount of uniformly ^{15}N,^{13}C-labelled native Curli from the *Escherichia coli* MC4100 strain for solid-state NMR analysis. A 2D ^{13}C-^{13}C solid-state NMR spectrum of native Curli revealed a striking similarity to the spectra obtained from recombinant CsgA fibrils, implying that CsgA fibrils adopt a highly similar tertiary structure compared to native Curli.

CsgB can nucleate CsgA to form mixed CsgA-CsgB fibrils with any mixing ratio. The mixed fibrils revealed a highly ordered, amyloid-like conformation similar to native Curli. I investigated *in-vitro* fibrilized mixed fibrils using immunogold labelling TEM, controlled fragmentation by shear-force combined with seeding assays and H/D exchange NMR. Immungold labelling TEM images of mixed fibrils clearly displayed that CsgB is located not only at the ends but also in between CsgA molecules in mixed fibrils. Both the N-terminal region and C-terminal region of CsgB could interact with neighbouring CsgA molecules respectively. Mixed CsgA-CsgB fibrils appeared to be fragmented into optimal seeds at slightly lower shear forces compared to pure CsgA fibrils, possibly indicating that interspersed CsgB weakens the overall stability of CsgA fibrils. H/D exchange data showed that interspersed CsgB formed a more ordered β-sheet structure in mixed fibrils.

1. Introduction

1.1. Amyloid

Amylum is the Latin word for the normal plant amylaceous material, which was firstly described in 1838 by the German botanist Matthias Jakob Schleiden (Kyle 2001). In 1854, Rudolf Virchow firstly used the term "amyloid" to describe the abnormal extracellular deposits that are seen in the liver during autopsy that exhibited a positive iodine staining reaction (reviewed in (Sipe and Cohen 2000)). In 1927, Divry found that the amyloid displayed apple-green birefringence when treated with Congo red dye and viewed under cross-polarized light (Sipe and Cohen 2000). This is one of the most important approaches which is until now used to identify amyloid in tissue samples. In 1959, with the use of electron microscopy, Cohen discovered that the amyloid demonstrated a non-branching fibrillar structure (Cohen and Calkins 1959).

Amyloids are now defined as long, non-branching fibrous protein aggregates sharing β-sheets structural feature, and which are assembled through the templated homotypic polymerization of up to hundreds of thousands of monomeric polypeptides (Kyle 2001). The structure of amyloid fibrils of HET-s (218-289) was investigated using solid-state NMR in 2005 (Ritter, Maddelein et al. 2005). In the last ten years, solid-state nuclear magnetic resonance (ssNMR) spectrascopy has become the most powerful tool to study atomic-level structure of amyloid proteins (Luhrs, Ritter et al. 2005, Ritter, Maddelein et al. 2005, Iwata, Fujiwara et al. 2006, Wasmer, Lange et al. 2008, Nielsen, Bjerring et al. 2009).

1.1.1. Amyloids and Amyloidosis

Amyloidosis is a rare disease that results from accumulation of inappropriately folded proteins-amyloids. These amyloids accumulate in body tissue and organs and cause altered function or cells damage. The exact cause of amyloidosis is often unknown. However, there are several possible reasons. Most importantly, amyloids are associated with a variety of aging-associated degenerative diseases. These amyloids may be caused by the free radicals which are the harmful byproducts formed when cells use energy (Hensley, Carney et al. 1994, Sultana, Mecocci et al. 2011, Wilkinson, Cramer et al. 2012, Giliano, Stepanov et al. 2013, Palmer, Tayler et al. 2013). For example, when people get older, β-amyloid peeptides could

aggregate into β-amyloid aggregates and fibrils in the brain resulting in Alzheimer's disease. Amyloid could be also caused in people who have immune system damage or defects (Sultana, Mecocci et al. 2011, Mollica, Stefanucci et al. 2012). Familial amyloidosis resulting from genetic changes also causes the body to produce mutant amyloid proteins (Das, Mei et al. 2014, Martin, Benson et al. 2014, Yin, Xia et al. 2014).

Amyloidosis is classified as follows (Hazenberg 2013):

1) Primary amyloidosis, usually with no evidence of preceding or coexisting disease, plasma-cell dyscrasia or paraproteinemia;

2) Secondary amyloidosis with evidence of coexisting previous chronic inflammatory or infectious conditions;

3) Other amyloidosis that associate with multiple myeloma.

In amyloidosis, once amyloid start, they seem to continue building up in the same locations in body organs. The heart, kidneys, liver, nervous system, spleen and gastrointestinal tract are the most frequently affected by amyloidosis. At least 18 amyloids have been identified as well as implicated in the pathology of various human diseases (**Table 1**) (Chiti and Dobson 2006, Fowler, Koulov et al. 2007, Harrison, Sharpe et al. 2007, Herczenik and Gebbink 2008).

Table 1: Amyloids are associated with various diseases.

(Reviewed by (Chiti and Dobson 2006, Harrison, Sharpe et al. 2007, Herczenik and Gebbink 2008)

Amyloid	**Disease/ Symptom**
β-amyloid (Aβ)	Alzheimer's disease
α-synuclein	Parkinson's disease
Islet amyloid polypeptide (AIAPP)	Type 2 Diabetes
Serum amyloid A (SAA)	Rheumatoid arthritis
β (2) microglobulin amyloid (Aβ2M)	Dialysis related amyloidosis
Prion protein (PrPc)	Spongiform encephalopathies
Huntingtin	Huntington's Disease
Calcitonin	Medullary carcinoma of the thyroid
Atrial amyloid	Cardiac arrhythmias, Isolated atrial amyloidosis
Apolipoprotein AI (AApoA1)	Atherosclerosis
Medin amyloid (AMed)	Aortic medial amyloidosis
Prolactin amyloid (APro)	Prolactinomas
Transthyretin amyloid (ATTR)	Familial amyloid polyneuropathy
Lysozyme	Hereditary non-neuropathic systemic amyloidosis
Gelsolin	Finnish amyloidosis
Keratoepithelin	Lattice corneal dystrophy
Cystatin	Congophilic angiopathy
Immunoglobulin light chain AL	Systemic AL amyloidosis

1.1.2. Functional amyloids

Functional amyloid is proteinaceous fibers forming in physiological conditions that are found in bacteria, fungi, invertebrates, and not least, humans (Gebbink, Claessen et al. 2005). Unlike pathogenic amyloids, such well-ordered functional amyloid have nonpathogenic properties crucial e.g. to biofilm formation and virulence (Otzen and Nielsen 2008). Within the past decade, several functional amyloids have been identified on the surfaces of bacteria and fungi (Gebbink, Claessen et al. 2005, Fowler, Koulov et al. 2006, Otzen and Nielsen 2008, Greenwald and Riek 2010) (**Table 2**). For example, the protein CsgA form Curli amyloid on the cell surfaces of *Escherichia coli* (Olsen, Jonsson et al. 1989, Chapman, Robinson et al. 2002); and hydrophobins are the proteins known to form amyloids on fungal surfaces (Wosten 2001).

Table 2: Functional amyloids

(Reviewed by (Gebbink, Claessen et al. 2005, Otzen and Nielsen 2008))
A summary of 16 amyloid proteins is coming from different organisms, which have many biological functions.

Amyloid	Organism	Function
Curli	*Escherichia coli and other Enterobacteriaceae*	Cell adhesion and invasion, biofilm formation, attachment to host surface
MTP	*Mycobacterium tuberculosis*	Adhesive properties on pili
Chaplins	*Streptomyces coelicolor*	Formation of aerial hyphae and modulation of water surface tension
FapC	*Pseudomonas fluorescens*	Fimbriae formation
Microcin E492	*Klebsiella pneumonia*	Cytotoxin, forms pores in cytoplasmic membrane
TasA	*Bacillus subtilis*	Biofilm formation
Sup35p	*Saccharomyces cerevisiae*	Translation termination
Ure2p	*Saccharomyces cerevisiae*	Native state involved in nitrogen catabolism
Rnq1p	*Saccharomyces cerevisiae*	Chromatin remodeling
HET-s	*Podospora anserina*	Regulation of heterokaryon formation
CPEB	*Aplysia californica*	Promotion of long-term maintenance of synaptic changes associated with memory storage
HpaG	*Xanthomonas spp.and other plant pathogens*	Cytotoxin, hairpin causes hypersensitive response in plant cells
Spidroins	*Nephila edulis*	Silk fibers of the web
Silkmothchorion	*Bombyx mori*	Protection of the oocyte and developing embryo
Hydrophobins	*Neurospora crassa and other fungi*	Modulation of surface attachment and aerial hyphae formation
Mα (Pmel17)	*Human*	Melanin biosynthesis

1.1.3. Amyloid folds

The primary structural trait of amyloid fibrils is a characteristic β-sheet structure termed cross-β structure. Amyloid fibrils consist of well-aligned polypeptides that form β-sheets that

run parallel to the fiber axis, with β-strands perpendicular to the fiber axis. The β-strands can be arranged in either a parallel or an antiparallel ortientation (Toyama and Weissman 2011).
The sheet structures are stabilized by non-covalent intramolecular and intermolecular interactions, such as aromatic stacking interactions and hydrogen bonds. Briefly, individual β-sheets interact with neighbor β-sheets by two types of β-sheet stacking interfaces that are termed the dry and wet interfaces. In dry interfaces, the well-packed side chains interlink by aromatic stacking interactions and H-bonds resulted in close proximity, this is termed the steric zipper (Greenwald and Riek 2010). The wet interface is lined with water molecules that make the wet interface wet (Nelson, Sawaya et al. 2005, Greenwald and Riek 2010).

1.1.4. Amyloid biophysics

Structural studies on amyloids started with the observation of globulins using X-ray diffraction by Astbury and Dickinson in 1935 (Astbury, Dickinson et al. 1935). X-ray fiber diffraction demonstrated that the inter-strand and inter-sheet stacking distances of the cross-β core are 4.7Å and 10Å respectively (Astbury, Dickinson et al. 1935, Sunde, Serpell et al. 1997).
Because of the insoluble property of amyloid fibrils, the common biophysical techniques such as X-ray crystallography and solution nuclear magnetic resonance (solution NMR) cannot be straightforward used to reveal the structural informations of amyloid fibrils. However, over the past decades, structural biology in amyloid fibrils has made considerable progress, the biophysical approaches have enabled higher-resolution insights into the atomic details of several amyloid fibrils, such as X-ray fiber diffraction (Astbury, Dickinson et al. 1935, Sunde, Serpell et al. 1997), Cryoelectron microscopy (cryoEM) (Jimenez, Guijarro et al. 1999, Jimenez, Nettleton et al. 2002, Zhang, Hu et al. 2009), Hydrogen/Deuterium exchange NMR (Hoshino, Katou et al. 2002, Luhrs, Ritter et al. 2005, Ritter, Maddelein et al. 2005, Vilar, Wang et al. 2012) and solid-state NMR (Iwata, Fujiwara et al. 2006, Lim, Nguyen et al. 2006, Baxa, Wickner et al. 2007, Wasmer, Lange et al. 2008, Shewmaker, McGlinchey et al. 2009, Wasmer, Benkemoun et al. 2009, Scheidt, Morgado et al. 2012).

1.1.5. Amyloid toxicity and clearance

Mature fibrils or aggragate fibrils have long been considered as the cause of disease; however, in recent years, oligomers or prefibrillar aggregates have been shown significant cytotoxicity

in neurodegenerative diseases (Ono, Condron et al. 2009). For example, both oligomeric Aβ and prefibrillar Aβ aggregates are potent neurotoxins resulting in synaptic loss and severity of cognitive impairment (Pike, Walencewicz et al. 1991, Townsend, Shankar et al. 2006, Li, Dolios et al. 2014).

The formation of oligomers or prefibrillar aggregates could cause neurodegenerative diseases; that is to say, the clearance of oligomers or prefibrillar aggregates could be a possible strategy for neurodegenerative diseases treatment. For example, the metalloproteinase of insulin-degrading enzyme (IDE) and endothelin-converting enzyme (ECE) have been discovered that selectively degraded monomers or aggregates and contributed to clearance of Aβ aggregates in the brain (Eckman, Watson et al. 2003, Palmer, Baig et al. 2009, de Tullio, Castelletto et al. 2013, McCord, Liang et al. 2013).

1.2.*Escherichia coli* amyloid Curli

1.2.1. Functions of Curli

Extracellular functional amyloids of many enteric bacteria including *E.coli* and Salmonella species are crucial for pathogenesis. In the pathogenesis of infection, The Curli of *E.coli* is involved in the attachment and invasion of host cells and activation of the immune system (Barnhart and Chapman 2006) (**Table 3**).

Curli can directly interact with many host matrix proteins such as plasminogen and its physiological activator-tissue plasminogen activator (t-PA), these proteins are proposed to stimulate bacterial dissemination through the host. Plasminogen is a glycoprotein that is composed of an inactive serine protease domain and five homologous triple-loop domains (also termed ‘kringles’) (Sjobring, Pohl et al. 1994, Sipe and Cohen 2000, Castellino and Ploplis 2005). Plasminogen presents in plasma and extracellular fluids and can degrade fibrin and soft tissue (Castellino and Ploplis 2005). t-PA is a serine protease which can activate plasminogen to plasmin (Robbins, Summaria et al. 1967). Curliated bacteria were able to absorb plasminogen from eukaryotic cell plasma, and further the captured plasminogen can be readily activated by bound t-PA. This pathway could be related to induce destruction of tissue barriers and empower the bacteria to gain access to deeper tissues (Saksela and Rifkin 1988, Sjobring, Pohl et al. 1994).

E.coli Curli can also mediate internalization of bacteria by eukaryotic cells. Curli bind to the extracellular-matrix soluble fibronectin protein with high affinity strongly, resulting in Curli-

mediated internalization (Olsen, Jonsson et al. 1989, Collinson, Doig et al. 1993, Gophna, Oelschlaeger et al. 2002).

Table 3: Proteins that interact with amyloid Curli

A summary of 9 proteins can interact with amyloid Curli showing different biological functions.

Protein	Function
Plasminogen	Extracellular matrix protein
Tissue plasminogen activator (t-PA)	Converts plasminogen to plasmin
Fibronectin	Degrade fibrin and soft tissue
MHC class 1	Antigen-presenting molecule
Toll-like receptor 1&2	Innate immune response
Fibrinogen	Human contact-phase protein
Coagulation factor XI	Human contact-phase protein
Coagulation factor XII	Human contact-phase protein
H-kininogen	Human contact-phase protein

Curli also binds to the proteins of the immune system. Major histocompatibility complex (MHC) class I (MHC-I) is a highly polymorphic transmembrane glycoproteins complex, which is composed of a 40 kDa heavy chain and a 12 kDa light chain (β2-microglobulin (β2m)) (Bjorkman, Saper et al. 1987). The receptor of MHC-I can present foreign peptides or antigens to cytolytic T cells and the peptides presentation by MHC-I is a central feature of anti-microbial immunity (Germain and Margulies 1993, Yewdell, Norbury et al. 1999). Both purified Curli and CsgA protein were recognized by intact monomeric MHC-I molecules and β2m (Olsen, Wick et al. 1998). Curliated bacteria also recognized by the MHC-I expressing macrophage cells (Olsen, Wick et al. 1998, Johansson, Nilsson et al. 2001), However, the curliated bacteria didn't significantly contribute to bacterial antigens processing and presentation (Johansson, Nilsson et al. 2001).

Toll-like receptors (TLRs) are pattern recognition receptors which recognize conserved molecular patterns of microbes as well as endogenous danger molecules as the pathogen-associated molecular pattern (PAMP) initiating inflammatory immune responses (Takeuchi and Akira 2010). Curli has been shown to be a PAMP molecule that is recognized by the TLR1/TLR2 complex (Tukel, Raffatellu et al. 2005, Tukel, Nishimori et al. 2010). Curli and recombinant CsgA protein have been also discovered that they can bind to CD14 protein respectively resulting in enhancing the activation of NF-κB through the TLR2/TLR1 complex

(Rapsinski, Newman et al. 2013), In short, Curli fibrils, released from the bacterial biofilm or still attached to bacteria, are recognized by the CD14-TLR2-TLR1 complex to trigger activation of NF-κB, and further NF-κB can induce higher levels of proinflammatory cytokines such as IL-2, IL-6 and IL-8 (Bian, Brauner et al. 2000, Tukel, Raffatellu et al. 2005, Tukel, Wilson et al. 2009, Rapsinski, Newman et al. 2013).

Curli also interacts with the human contact-phase factors such as coagulation factor XI and XII, fibrinogen and H-kininogen (Ben Nasr, Olsen et al. 1996, Herwald, Morgelin et al. 1998, Olsen, Herwald et al. 2002, Wang, Zhou et al. 2010). Curliated bacteria incubated with factor XI, factor XII and fibrinogen respectively resulting in the significantly decreasing of factor XII and fibrinogen production in supernatant (Herwald, Morgelin et al. 1998). Curliated bacteria can also prolong the clot formation by a deficiency of contact-phase factors, which could enhance the spread of bacteria to surrounding tissue (Herwald, Morgelin et al. 1998).

1.2.2. Curli and diseases

Amyloid Curli is involved in the pathogenesis of urinary tract infections. Urinary tract infections (UTIs) are common, ranking in prevalence with gastrointestinal and respiratory infections. Uropathogenic *E.coli* (UPEC) is the major causative agent of urinary tract infections. UPEC can invade into superficial cells of the bladder and then create pod-like bulges of complex intracellular bacterial communities (IBC) within the superficial umbrella cells of the bladder (Anderson, Palermo et al. 2003, Justice, Hung et al. 2004). The role of bacterial biofilm on urinary tract infections has been established, and bacterial biofilms on urinary catheters result in persistent infections that are resistant to antimicrobial factors (Trautner and Darouiche 2004). Curli can protect the bacteria from being killed by antimicrobial peptide LL-37, because LL-37 inhibits Curli assembly by preventing the polymerization of the major Curli subunit-CsgA (Chromek, Slamova et al. 2006, Wang, Smith et al. 2007, Kai-Larsen, Luthje et al. 2010). UPEC isolates collected from urine of UTI patients that had a higher adhesion capacity and produced more biofilm than commensal *E. coli* (Kai-Larsen, Luthje et al. 2010). Deletion of the csgA gene in a prototypic UPEC resulted in decreasing bladder colonizations at 6 hours post-infection in UTI model mice. These results exhibited that Curli plays an important role in bladder colonization (Cegelski, Pinkner et al. 2009).

Amyloid Curli also cause urinary source bloodstream infection. Curli fibrils are known to be expressed at temperatures below 26°C, however, *E. coli* isolates from hospitalized patients

showed robust Curli production at physiologic temperature of 37°C (Olsen, Arnqvist et al. 1993). This result indicated that Curli expression at 37°C by urinary tract *E.coli* strains was associated with bacteremic progression. The ERIC (Enterobacterial Repetitive Intergenic Consensus)-PCR analysis of same-patients urine and blood isolates exhibited a high correlation. These findings suggested that urinary tract *E. coli* could directly invade into the bloodstream (Hung, Marschall et al. 2014).

1.2.3. Curli biogenesis

Two adjacent divergently transcribed operons, CsgBAC and CsgDEFG, express 7 Csg proteins in all, termed CsgA, CsgB, CsgC, CsgD, CsgE, CsgF and CsgG (Chapman, Robinson et al. 2002, Brombacher, Baratto et al. 2006). CsgA and CsgB are two structural proteins that are involved in Curli assembly (Chapman, Robinson et al. 2002). CsgC, CsgE, CsgF and CsgG are non-structural proteins that are involved in stability and secretion of CsgA CsgB and CsgF subunits (Robinson, Ashman et al. 2006, Nenninger, Robinson et al. 2009, Taylor, Zhou et al. 2011). CsgD, a transcriptional activator, regulates the expression of 6 Csg proteins under environmental conditions such as osmolality and temperature (Brombacher, Baratto et al. 2006) (**Figure 1**).

Curli is composed of a major protein CsgA and a minor nucleator protein CsgB (Olsen, Jonsson et al. 1989, Chapman, Robinson et al. 2002, Chiti and Dobson 2006). During Curli assembly, the soluble and unfolded CsgB monomers are initially secreted to the outer membrane surface across the CsgG pore. CsgA monomers are also secreted in the same manner as unfolded proteins to the bacterial surface. On the cell surface, the unfolded CsgA monomer folds into the proper conformation on association with folded CsgB, which act as the nucleation center. Sequent new unfolded CsgA monomers immediately present a proper conformation on association with the nucleation centers resulting in well-ordered amyloid fibers (Hammar, Bian et al. 1996, Hung, Marschall et al. 2014). In CsgB mutant bacteria, CsgA is secreted to the extracelluar space but cannot polymerize into amyloid fibers. In the absence of CsgA, bacteria don't produce amyloid Curli either. However, bacteria continued to express CsgB protein after 48 h of growth (Wang, Hammer et al. 2008).

In vitro, freshly purified CsgA and CsgB proteins exhibit intrinsically disordered properties as monomers. Both CsgA and CsgB individually are able to self-aggregate to form fibrils under proper conditions. The speed of CsgB self-aggregation is faster than the speed of CsgA self-aggregation. CsgB and $CsgB^{\Delta R5}$ were proved to nucleate unstructured CsgA into fiber *in vitro*,

which might suggest that CsgB could nucleate *in-vivo* CsgA to form amyloid fibers (Hammer, Schmidt et al. 2007, Wang, Hammer et al. 2008).

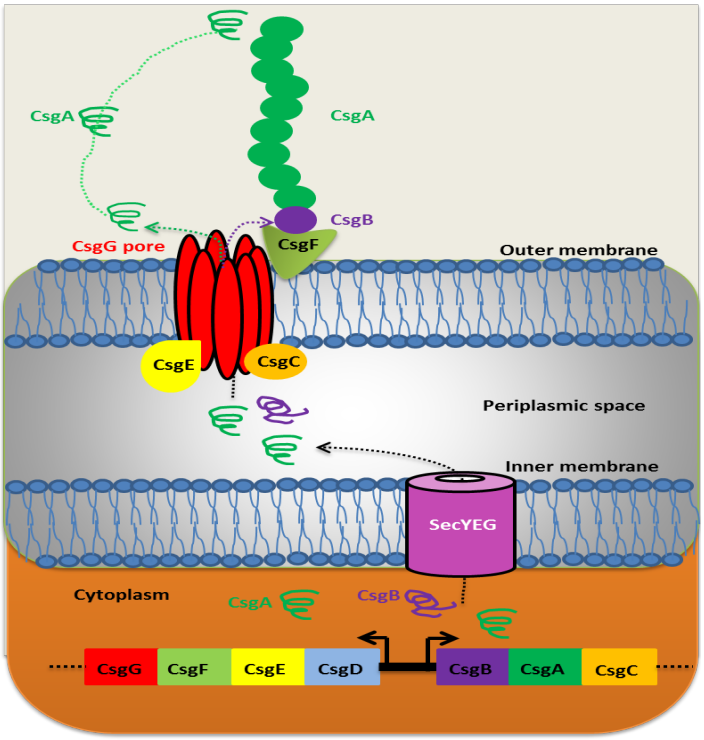


Figure 1: *E.coli* Curli biogenesis

The csg genes involved in *E.coli* amyloid Curli assembly are organized into two adjacent divergently transcribed operons, csgBAC and csgDEFG, CsgA and CsgB are two structural subunits which are involved in Curli assembly. CsgC, CsgE, CsgF and CsgG are non-structural proteins that are involved in stability and secretion of CsgA and CsgB subunits. CsgD, a transcriptional activator, regulates all Csg proteins production.

CsgF is critical for CsgB localization on the bacterial surface. In absence of CsgF, CsgB failed to associate with the bacterial surface resulting in CsgA secretion away from bacterial cells, and CsgA polymerization that was both imcomplete and mislocalized. In presence of CsgF, CsgB was well-localized on the bacterial surface, and further formed well-ordered amyloid fibrils resistant to protease degradation compared with CsgB produced in the absence of CsgF. Taken these results together, CsgF determines where and how Curli assembling extracellularly (Nenninger, Robinson et al. 2009).

CsgC is the first Csg protein structure to be solved by X-ray crystallography. Crystal structure of CsgC adopts an immunoglobulin (Ig)-like β sandwich with seven strands; the structure is similar to the N-terminal domain of DsbD protein, an oxido-reductase superfamily CxxC-like

motif. The similarity of structure suggested that CsgC could be involved in redox activity in Curli biogenesis. Mutant strains lacking CsgC showed decreasing bile salt influx, but the presence of CsgC, increaseing the flux of CsgA and CsgB proteins though CsgG. The proposed mechanism is that CsgC modulates the activity of the CsgG pore activity though catalyzing disulfide-bonds formation to crosslink CsgG transmembrane helices (Taylor, Zhou et al. 2011).

CsgD, a transcriptional activator, is a homologous member of the FixJ/UhpA/LuxR family. CsgD regulates Curli and cellulose production in *E.coli* (Chirwa and Herrington 2003). Two antisence sRNAs, OmrA and OmrB, were overexpressed resulting in Curli deficiency. Further experiments showed that both OmrA and OmrB downregulated and inhibited CsgD translation *in vitro* (Holmqvist, Reimegard et al. 2010). The expression of CsgD was regulated by RpoS sigma factor as well though the positive action of the transcriptional regulator MlrA. The CsgD expression was also controlled by several global regulars such as IHF, CpxR, OmpR and H-NS depending on bacterial growth condition (Chirwa and Herrington 2003, Gerstel, Park et al. 2003, Brombacher, Baratto et al. 2006, Da Re and Ghigo 2006).

CsgE, proposed to be a chaperone-like protein, participates in a complex with CsgG protein. It might affect CsgG pore secretion activity. In CsgE mutant, there were less extracellular Curli fibers detected by EM, and also less CsgA and CsgB were detected by western-blot analysis (Nenninger, Robinson et al. 2011). Interestingly, CsgE can block unspecific CsgG mediated secretion (Nenninger, Robinson et al. 2011). In absence of CsgE, when overexpression of CsgG, two small perplasmic proteins PapD2 (93aa) and Cpxp (145aa) can be detected in the supernatant by western blot analysis. However, PapD2 and Cpxp weren't detected in the supernatant in the absence of CsgG (Robinson, Ashman et al. 2006). These results suggested that in absence of CsgE, the overexpressed CsgG pore could be ungated to translocate smaller (the protein size is close to Csg protein) perplasmic proteins and Csg proteins (Robinson, Ashman et al. 2006, Nenninger, Robinson et al. 2011).
CsgE can inhibit *in-vitro* CsgA aggregation. Freshly purified unfolded CsgA protein was incubated alone or with freshly purified CsgE in a near 1:1 molar ratio for 24 h, Alone, CsgA can self-aggregate into large collections of 4-6 nm CsgA fibrils, however, CsgA failed to assemble to fibrils when incubated with CsgE protein. ThT fluorescence assays also supported these results (Nenninger, Robinson et al. 2011).

Collectively, CsgE efficiently modulates the gating and ungating of CsgG pore to allow only CsgA, CsgB and CsgF secretion at physiological levels of CsgG expression as well as make Curli assembly more efficient (Nenninger, Robinson et al. 2011).

CsgG, a lipoprotein, localizes to the outer membrane forming an oligomeric annular shaped structure (approximately 2nm in diameter) (Robinson, Ashman et al. 2006, Taylor, Zhou et al. 2011, Goyal, Krasteva et al. 2014). The secretion and localization of CsgA, CsgB and CsgF proteins were dependent on CsgG pore (Epstein, Reizian et al. 2009). In absence of CsgG, CsgA, CsgB and CsgF weren't secreted to the cell surface. CsgA, CsgB and CsgF were also extremely unstable in absence of CsgG protein (Robinson, Ashman et al. 2006, Nenninger, Robinson et al. 2011, Taylor, Zhou et al. 2011).

The N22 region of CsgA doesn't participate in the amyloid core. The N22 was fused with Papd2 to create N22-Papd2 fusion protein in LSR11 cells, in presence of CsgG, N22-Papd2 can secrete to bacterial surface, and can be detected by western blot analysis. However, no N22-Papd2 fusion protein was detected in bacterial surface in absence of CsgG. These data could suggest that the N22 is the specificity signal to allow Csg proteins across the CsgG pore (Nenninger, Robinson et al. 2011).

1.2.4. Sequence analysis of CsgA and CsgB subunits

(1) Sequence analysis of CsgA subunit

CsgA, the major subunit of amyloid Curli, can be divided into three identifiable domains. The N-terminus contains a sec-dependent signal (residues 1-20) segment that is cleaved after translocation across the cytoplasmic membrane (Collinson, Parker et al. 1999). The N-terminal 22 amino acid segment (residues 21-42, termed N22) doesn't constitute an integral part of CsgA fibrils, and the N22 domain was proposed to directly bind to CsgG resulting in CsgA secretion and stability (Robinson, Ashman et al. 2006, Wang, Smith et al. 2007). The C-terminal amyloid core segment (residues 43-151) is composed of five imperfect repeating units (R1-R5). Each repeat (19-23 residues) is distinguished by the consensus sequence of Ser-X5-Gln-X5-Asn-X5-Gln (X represents any residue) (Chiti and Dobson 2006, Wang, Smith et al. 2007). These Ser, Gln and Asn residues are extremely conserved among CsgA homologs from all known entericbacteria, and are crucial for Curli formation (Hammer, Schmidt et al. 2007, Wang, Smith et al. 2007) (Figure 2A, Figure 3C).

The sequences of R1 and R5 are crucial for CsgA self-aggregation and CsgB nucleation

of CsgA. Each repeat is suggested to form a β strand-loop- β strand motif. Synthesized peptides R1, R3 and R5 have been found to efficiently assemble into amyloid-like fibrils *in vitro*, while R2 and R4 didn't (Wang, Smith et al. 2007). Furthermore, WT CsgA and five CsgA mutants- $CsgA^{\Delta R1}$, $CsgA^{\Delta R2}$, $CsgA^{\Delta R3}$, $CsgA^{\Delta R4}$, $CsgA^{\Delta R5}$ were freshly purified to test the capability of these mutants to self-aggregate *in vitro* using the ThT fluorescence assay, the lag phase and fiber growth phase for $CsgA^{\Delta R1}$, $CsgA^{\Delta R2}$, $CsgA^{\Delta R3}$ and $CsgA^{\Delta R4}$ mutants were quite similar to those of WT CsgA protein, however, both lag phase and growth phase for $CsgA^{\Delta R5}$ mutant were significantly increased (Wang and Chapman 2008).

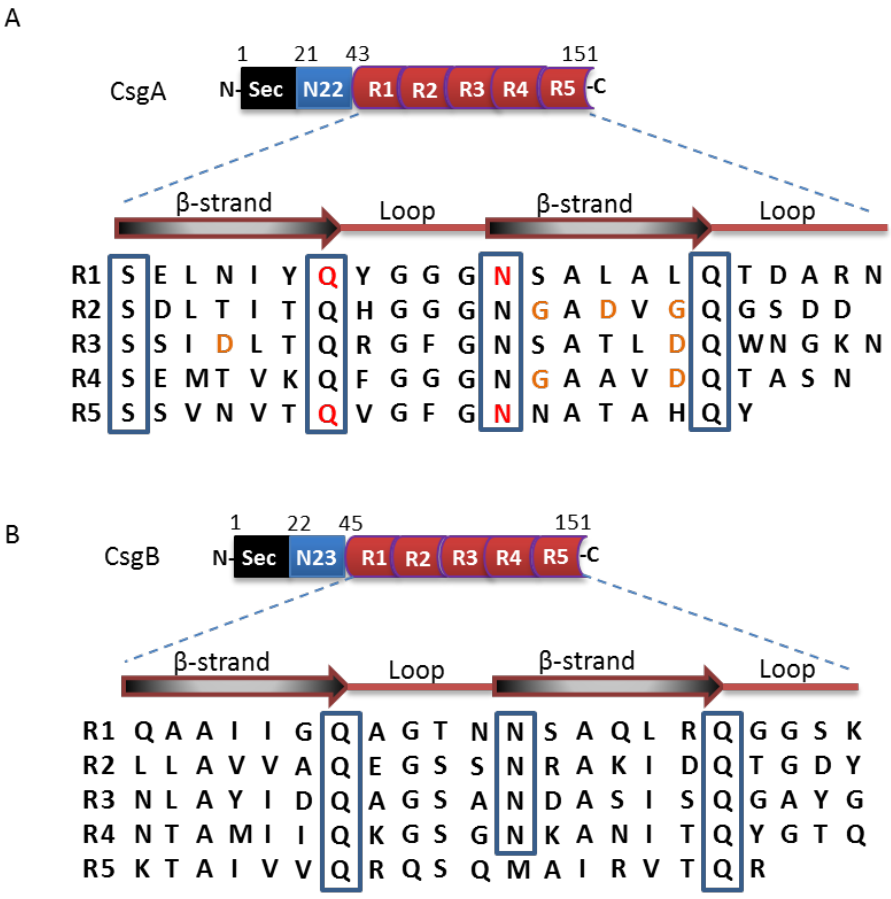


Figure 2: Protein Sequences of CsgA and CsgB subunits

Both CsgA (A) and CsgB (B) are composed of three domains, the first domain is the N terminal sec signal domain; the second domain is the N-terminal 22 aa peptide domain of CsgA and the N-terminal 23 aa peptide domain of CsgB, the third domain is the amyloid core which is composed of five imperfect repeating units (R1-R5). (A) Selected four conserved Q49, N54, Q139 and N144 residues are shown in red. Selected gatekeeper residues G78, D80, G82, D91, D104, G123 and D127 are shown in yellow.

Conserved Gln(Q) and Aln(N) residues of R1 and R5 repeats contribute significantly to Curli assembly (Wang, Smith et al. 2007, Wang, Hammer et al. 2008). An Ala scanning mutagenesis experiment of the 20 conserved residues was performed to test these residues

importance in bacterial amyloid formation in YESCA agar plates containing congo red, the results showed that Q49A and N144A mutants dramatically reduced congo red binding compared with WT CsgA (Wang, Hammer et al. 2008). The protein expression in these mutants was detected by Western blot analysis, the results revealed that the mutant proteins of CsgAQ49A, CsgAN54A, CsgAQ139A and CsgAN144A can be secreted to the bacterial surface, but they were significantly defective in Curli assembly. In addition, freshly purified recombinant CsgAQ49A and CsgAN144A mutants were also defective in self-aggregation in vitro. However, when in presence of WT CsgA fibrils, freshly purified CsgAQ49A and CsgAN144A mutants efficiently aggregated similar to that of WT CsgA protein. More interestingly, freshly purified CsgAQ49A, CsgAN54A, CsgAQ139A and CsgAN144A mutants were defective in amyloid fibers formation in CsgB-presenting cell mutant (csgA-csgB+) or freshly purified nuceator CsgB and CsgB truncation mutant (CsgBΔR5) (Wang, Hammer et al. 2008).

Peptides R2, R3 and R4 have poor amyloidogenic properties. These peptides contain "gatekeeper" residues which can slow CsgA aggregation (Wang, Hammer et al. 2008, Wang, Zhou et al. 2010). The 7 residues, Gly^{78}, Asp^{80} and Gly^{82} within R2, Asp^{91} and Asp^{104} within R3, Gly^{123} and Asp^{127} within R4, were identified as the "gatekeeper" respectively. Mutation of these "gatekeeper" residues affected CsgA or peptide polymerization (Wang, Zhou et al. 2010).

$CsgA^{*}$, a CsgA mutant lacking all gatekeeper residues ($CsgA^{G78S/D80L/G82L/D91N/D104L/G123N/D127H}$) rapidly self-aggragated into amyloid *in vitro* and did not even have a lag phase. Interestingly, $CsgA^{*}$ was efficiently polymerized *in vivo* in absence of CsgB and CsgF. These gatekeeper residues might modulate a CsgB/CsgF-dependent CsgA assembly as well as is mostly unattached to the bacteria (Wang, Zhou et al. 2010).

(2) Sequence analysis of CsgB subunit

CsgB, the minor subunit of Curli, can be also divided into three domains. a sec signal domain (residues 1-21), a N23 domain (residues 22-44) and a C-terminal amyloid core domain (residues 45-151) is composed of five repeating units (R1-R5, 19-22 residues for each unit)(Hammar, Bian et al. 1996, Bian and Normark 1997) **(Figure 2B)**. The consensus sequence of the R1, R2, R3 and R4 are X6-Gln-X5-Asn-X5-Gln containing highly conserved residues Gln and Asn of each repeating unit. The R5 contains four positively charged amino acids (Lys133, Arg140, Arg147 and Arg151) (Bian and Normark 1997, White, Collinson et al. 2001, Hammer, Schmidt et al. 2007).

Curli assembly on the bacterial surface depends on the nucleator CsgB protein. CsgB shares approximately 49% sequence identity with CsgA (Hammar, Bian et al. 1996). Structural prediction suggested that CsgA and CsgB might share biochemical properties (Collinson, Parker et al. 1999, White, Collinson et al. 2001). Both purified WT CsgB and $CsgB^{\Delta R5}$ can self-aggregate into amyloid-like fibrils. At equal concentrations, the lag phase of WT CsgB (20-30min) was shorter than $CsgB^{\Delta R5}$ (approximately 120min) (Hammer, Schmidt et al. 2007, Hammer, McGuffie et al. 2012). The $CsgB^{\Delta R5}$ aggregates can efficiently shorten the lag phase of CsgA polymerization (Hammer, Schmidt et al. 2007), this specific process is nucleation. $CsgB^{\Delta r5}$ can mediate CsgA nucleation *in vivo* (Hammer, McGuffie et al. 2012). WT CsgB and $CsgB^{\Delta r5}$ were expressed in the Δcsg strain LSR12 that contained CsgG respectively, both WT CsgB and $CsgB^{\Delta r5}$ proteins were expressed in SDS-soluble forms (Hammer, Schmidt et al. 2007). WT CsgB was found on the cell surface detected by western blot analysis; however, $CsgB^{\Delta r5}$ was expressed and secreted away from the bacterial surface into extracellular space. It assumed that R5 plays an important role in the cell membrane association (Hammer, Schmidt et al. 2007). WT CsgB was defective in cell surface association at an early stage of growth in the absence of CsgF *in vivo* (Nenninger, Robinson et al. 2009). However, How CsgB localization on the bacterial surface isn't still fully understood.

1.2.5. Biophysical characteristics of Curli fibrils

Both *E.coli* Curli (Figure 3A) and recombinant CsgA (Figure 3B) exbibit amyloid-like fibrils. Amyloid Curli and CsgA fibrils were extremely stable even in a few percent of SDS buffer as well as 5-8 M urea (Hammar, Bian et al. 1996, Chapman, Robinson et al. 2002, Shewmaker, McGlinchey et al. 2009). Incredibly, Curli can't be denatured after boiling for 10 minutes in SDS-PAGE loading buffer (Hammar, Bian et al. 1996). CsgA fibrils exhibited high levels of proteinase K resistance relative to soluble CsgA protein (Shewmaker, McGlinchey et al. 2009). But both Curli and CsgA fibrils can efficiently dissolve in 90% formic acid (Hammar, Bian et al. 1996, Chapman, Robinson et al. 2002).

CsgA fibrils exhibited strong β-sheet signatures (218nm) in CD spectrum. The FT-IR spectra of CsgA fibrils showed a well-defined peak at 1623 cm^{-1} (amide I band) that also indicated amyloid specific β-sheet secondary structure (Kong and Yu 2007, Shewmaker, McGlinchey et al. 2009, Dueholm, Nielsen et al. 2011).

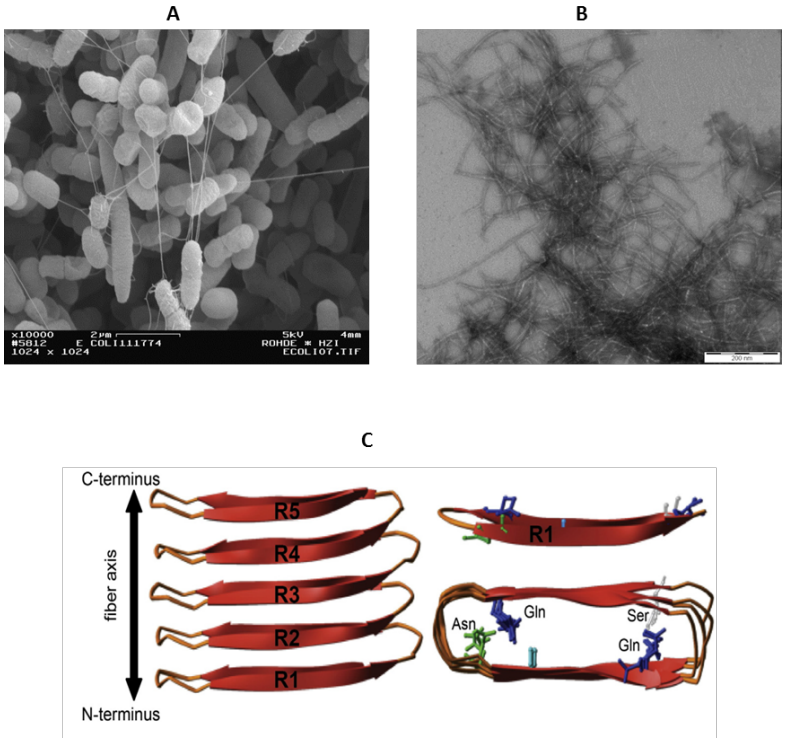


Figure 3: The predicted structure of Curli.

Negative-stained TEM micrographs of *E.coli* MC4100 ((A), scale bar represents 2 µm, took by Dr. Manfred Rohde and Dr. Madhu Nagaraj) and *in vitro* self-aggregated CsgA ((B), scale bar represents 200nm). C: A cartoon representation of the predicted structure of CsgA shows the five repeating units (R1-R5) assemble into a β-helical cross-β structure (side view, left). The conserved Gln, Asn and Ser residues of R1 (upper right) are shown as sticks and the overhead view of CsgA (low right) shows the alignment of these residues. ((C) is adapted from (Olubiyi, Frenzel et al. 2013))

Curli and CsgA fibrils are insoluble and non-crystalline, prohibiting structural studies of Curli or CsgA fibrils by X-ray crystallography or solution NMR. However, solid-state NMR is a powerful biophysical tool that has been used to study amyloid fibrils such as Aβ and HET-s(Luhrs, Ritter et al. 2005, Ritter, Maddelein et al. 2005, Lange, Gattin et al. 2009, Wasmer, Benkemoun et al. 2009, Parthasarathy, Long et al. 2011, Taylor, Zhou et al. 2011, Scheidt, Morgado et al. 2012, Huber, Ovchinnikova et al. 2014). Unlike HET-s (218-289) and yeast Ure2p amyloids, CsgA and CsgB fibrils are not in-register parallel β-sheet structure (Ritter, Maddelein et al. 2005, Sabate, Baxa et al. 2007, Wang, Smith et al. 2007, Wasmer, Lange et al. 2008, Shewmaker, McGlinchey et al. 2009). For X-ray fiber diffraction, Well-aligned β-sheet fibrils show characteristic meridional reflections at 4.7-4.8Å and equatorial reflections at 10Å (Malinchik, Inouye et al. 1998, Fowler, Koulov et al. 2007, Wang, Smith et al. 2007). X-ray fiber diffraction of CsgA fibrils indicated atomic spacing of 4.6 or 4.8 Å and 9 or 10Å (Shewmaker, McGlinchey et al. 2009, Dueholm, Nielsen et al. 2011). Soild-state NMR experiments indicated that CsgA fibrils is not based on an in-register parallel β-sheet structure

(Shewmaker, McGlinchey et al. 2009).

1.2.6. Kinetics of CsgA fibrilliation

The transition of monomer CsgA to CsgA fibrils can be investigated by the amyloid-specific dye-ThT (Krebs, Bromley et al. 2005, Chiti and Dobson 2006). The self-aggregation of CsgA (10 μM) was characterized by approximately 2 h of lag phase in pH 7 phosphate buffer. The lag phase can be dramatically shortened by the addition of native Curli or CsgA or CsgB fibrils (Chapman, Robinson et al. 2002, Hammer, Schmidt et al. 2007, Wang, Smith et al. 2007, Dueholm, and Nielsen et al. 2011). The lag time efficiently decreased with an increasing seed concentration as well and the initial ThT value scaled linearly with seed concentration (Dueholm, Nielsen et al. 2011). **Figure 4A** shows that CsgA aggregation speed up by adding the seeds.

The lag phase of CsgA aggregation was highly depended on buffer pH as well as protein concentration (Dueholm, Nielsen et al. 2011). At higher pH conditions (above pH 7), fibrillation of CsgA is slower than in lower pH buffers (below pH 7). It was suggested that negatively charged residues (Gln and Asn) at higher pH values decrease the propensity of CsgA to assemble a fibrillation prone nucleuses whereas the positive charges present below the pH 7 have minor effects (Wang, Smith et al. 2007, Dueholm, Nielsen et al. 2011). Protein concentration is another important factor to affect CsgA aggregation. The lag phase of CsgA aggregation decreased approximately 2-fold over an 8-fold increase in protein concentration. But interestingly, the fiber growth phase remained essentially invariant (Dueholm, Nielsen et al. 2011).

The CsgA aggregation at pH 7 was analyzed using ThT (or ANS) binding, SDS-PAGE, CD, TEM and X-ray diffraction respectively (Wang, Smith et al. 2007, Shewmaker, McGlinchey et al. 2009, Dueholm, Nielsen et al. 2011). These results revealed that CsgA aggregation is initiated from unfolded CsgA monomer to folded monomer, these folded monomers later convert into thin needle-like protofibrils; the small protofibrils further assemble into hydrophobic mature fibrils, which subsequently associate into larger aggregates (**Figure 4B**) (Dueholm, Nielsen et al. 2011).

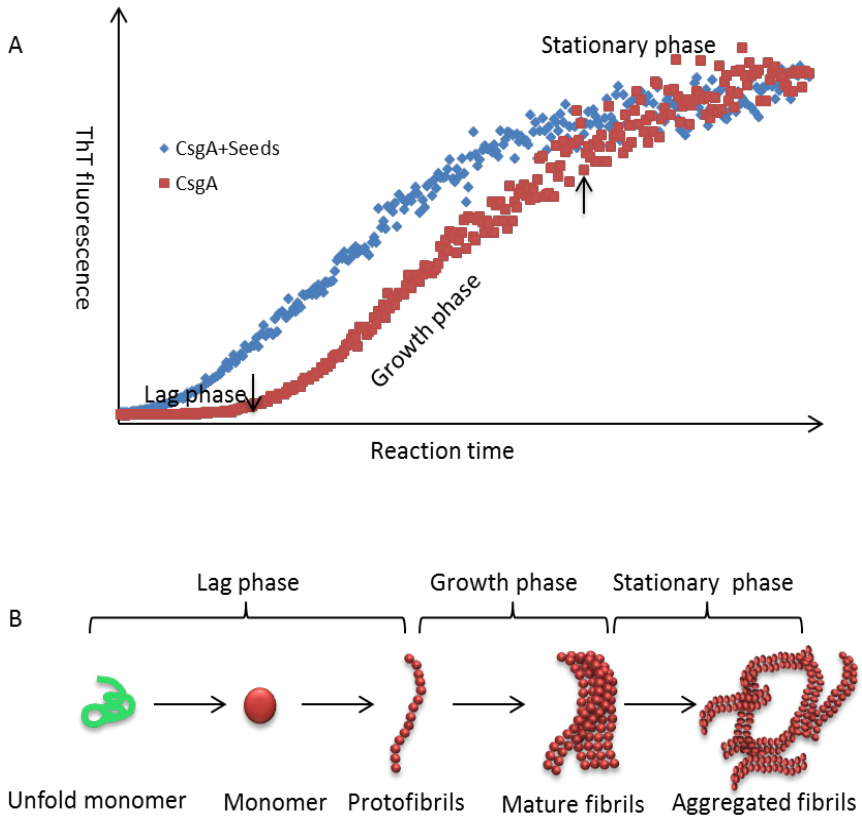


Figure 4: The proposed mechanism of CsgA fibrillation

A: CsgA self-aggregation is followed by ThT fluorescence. The curve of CsgA fibrils growth is composed of three steps-lag phase, growth phase and stationary phase (red curve). The lag phase can be efficiently shortened by the addition of amyloid fibrils as the seed (blue curve). B: During lag phase, unfold CsgA monomer forms folded monomer and further photofibrils. The growth phase starts at the formation of protofibrils; these protofibrils subsequently aggregate into aggregated firbils during stationary phase.

1.3. Techniques to study amyloid

1.3.1. Amyloid binding to cogno red and Thioflavin T dye

Amyloid fibrils are generally 5-16 nm in width and several micrometers in length and the fibrils have similar tinctorial properties, such as binding Congo red and Thioflavin T (ThT). Cogno red, a secondary diazo dye, binds to amyloid fibrils with a strongly non-covalent affinity. It can be used as a sensitive diagnosis tool for amyloidosis. Cogno red may bind in grooves formed along the β-sheets as a planar molecule of amyloid. (Khurana, Uversky et al. 2001, Howie and Brewer 2009, Schutz, Soragni et al. 2011).

ThT, a cationic benzothiazole dye, is used to identify amyloid proteins (Vassar and Culling 1959). In presence of amyloid fibrils, ThT shows an emission at 485 nm. However, in absence of amyloid fibrils, an emission occurs at 438 nm. Several proposed mechanisms were

explained how ThT recognizes and binds to amyloid fibrils, amyloid plaques in tissues or other β-rich peptides. The most accepted model was recently confirmed by experimental and theoretical work is that highly directional ThT monomer bind into β-sheet structured cavities with ThT long axis parallel to the elongation axis of the amyloid fibrils (Biancalana and Koide 2010).

1.3.2. Theoretical Background of NMR Spectroscopy

NMR is a physical phenomenon in which specific nuclei in a magnetic field absorb and re-emit electromagnetic radiation, which was first described and measured in molecular beams by Isidor Rabi in 1938 (Kolezhuk 1996). In recent decades, NMR spectroscopy is widely used and has proved to be the successfully method to reveal proteins structure at atomic resolution as well as elucidate proteins structure-function relationship under physiological conditions (Wuthrich 1990).

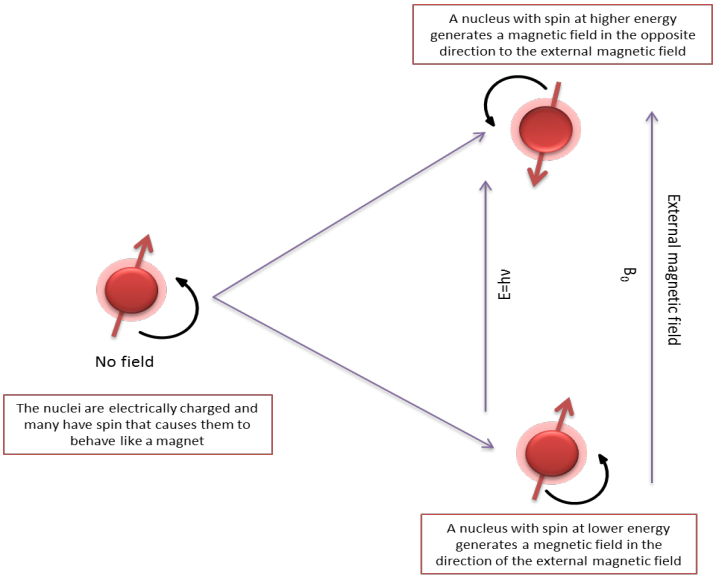


Figure 5: Splitting of nucleus state in an external magnetic field

^{1}H as well as many other nuclei such as ^{13}C and ^{15}N are related to spin 1/2 nuclei, the energy levels of these spin-½ nuclei can be split into two energy states when placed in external magnetic field.

NMR nucleus ^{1}H, and other nuclei such as ^{13}C and ^{15}N are used for NMR analysis, these nuclei have two possible spin states: $m = ½$ or $m = -½$. These states are degenerate without a magnetic, which is they have the same energy.

If a nucleus is placed in an external magnetic field, however, the interaction between the

nuclear magnetic moment (μ) and the external magnetic field (B_0) mean the two states no longer have the same energy. The energy of a magnetic moment when in a magnetic field is given by:

$$E = -\mu B_0$$

As a result the different nuclear spin states have different energies in a non-zero magnetic field. And this difference results in a small population bias toward the lower energy state.
The energy difference between the two states is:

$$\Delta E = 2\mu B_0$$

Resonant absorption by nuclear spins will occur only when electromagnetic radiation of the frequency is being applied to match the energy difference between the nuclear spin levels in external magnetic field. The energy of an absorbed photon is:

$$E = h\nu_0$$

(ν_0 is the radio frequency (RF)). Hence, a magnetic resonance absorption will only occur when

$$\Delta E = h\nu_0$$

Such magnetic resonance frequencies typically correspond to the RF range of the electromagnetic spectrum for magnetic fields. It is this magnetic resonant absorption that is detected in NMR spectroscopy **(Figure 5)**.

1.3.3. Hydrogen/Deuterium (H/D) exchange NMR spectroscopy

H/D exchange is a chemical reaction in which amide hydrogens in the peptide bonds of protein exchange protons with deuterons. The amide protons involved in the secondary structure display a slow exchange compared to the amide protons in the flexible regions of protein. Hence H/D exchange gives not only the chemical information about the solvent accessibility of various regions of protein, but also deeper reveals the presence of secondary structure in the protein (Ippel, Olofsson et al. 2002, Olofsson, Ippel et al. 2004, Jorgensen, Bache et al. 2005, Jorgensen, Gardsvoll et al. 2005).
Protons and deuterons can be easily differentiated by solution NMR spectroscopy. Standard HSQC spectra are recorded at specific time points after hydrogens in the buffer have been replaced by deuterium. The proton signal will decay exponentially as deuterium is invisible in

HSQC spectra. Afterwards, it is possible to fit an exponential function to the data, and get the exchange constant (Czerski, Vinogradova et al. 2000, Nabuurs and van Mierlo 2010, Vilar, Wang et al. 2012).

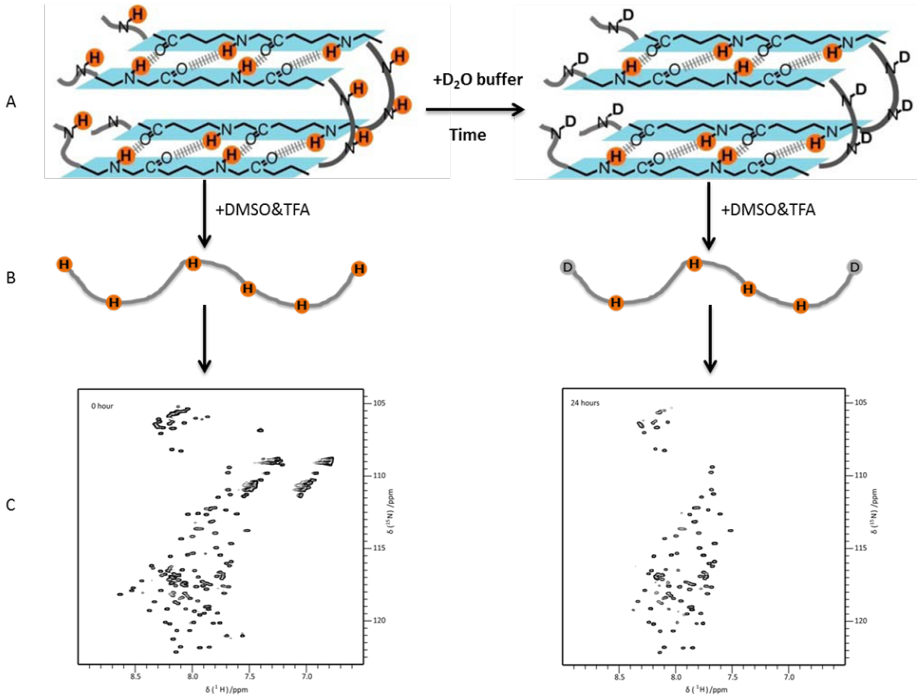


Figure 6: Schematic description of the H/D exchange NMR experiment used to study the secondary structure of amyloid fiber

(A is adapted from Christiane Ritter)
A: Protonated fibrils are incubated with D_2O buffer during the H/D exchange. Hydrogen bonded protons are protected from exchange, while the protons in the flexible region exchange to D_2O.
B: Monomerisation of the fibrils (A) using the proper solvent such as DMSO & TFA.
C: HSQC spectra are recorded at the specific time points (0 h and 24 h).

Biophysical methods for the determination of high-resolution structures of amyloid fibers are limited and challenging because amyloid fiber is insoluble as well as noncrystalline. However, H/D exchange NMR is a powerful biophysical tool used to directly investigate structural regions in amyloid fiber at single residue level (Englander 2000). The H/D exchange happens between the amide deuteriums of amyloid fiber and the protons of the buffer when amyloid fibers are incubated with the buffer (Olofsson, Lindhagen-Persson et al. 2009, Vilar, Wang et al. 2012). Importantly, the H/D exchange is significantly affected by the chemical environment of the amide deuteriums in the protein (Bai, Milne et al. 1993). The amide deuteriums in rigid regions are less accessible to the buffer because of the formation of deuterium bonds in the secondary structure, conversely, the amide deuteriums in flexible

regions easily exchange with protons (Dzwolak, Loksztejn et al. 2006, Olofsson, Lindhagen-Persson et al. 2009, Vilar, Wang et al. 2012). Exchanged amyloid fiber is dissolved in deuterated DMSO, and then records a HSQC spectrum of the exchanged sample (Dzwolak, Loksztejn et al. 2006, Vilar, Wang et al. 2012).

Generally, the H/D exchange NMR experiment of amyloid fibrils can be divided into three steps (Hoshino, Katou et al. 2002, Luhrs, Ritter et al. 2005), including：

(1) Exposure of the fibrils to deuterated buffer. The ^{15}N, ^{1}H-labelled fibrils are incubated with D_2O buffer in order to allow for hydrogen/deuterium exchange between the fibrils and D_2O buffer. The exchange reaction is quenched at specific time points by removing D_2O buffer, and then stored the exchanged fibrils in liquid nitrogen (**Figure 6A**).

(2) Rapid dissociation of the fibrils into random coli monomers using the proper solvent. The exchanged fibrils are rapidly dissociated into monomers using deuterated DMSO and TFA, TFA can promote the exchange due to residual water in the sample. The DMSO conserves the hydrogen/deuterium pattern present in the exchanged fibrils. Due to residual D_2O in the sample, unspecific H/D exchange will slowly take place for all residues. Therefore, this step is time-sensitive (Ritter, Maddelein et al. 2005) (**Figure 6B**).

(3) Detection of the exchanged sample by solution NMR spectroscopy.

^{15}N, ^{1}H-correlation spectra (HSQC) are recorded and analyzed for residual intensity of resonances in comparison to other spectra of the exchanged fibrils. The intersities are decreased with the increasing incubation times in D_2O buffer **(Figure 6C)**.

Another aspect is the analysis of NMR spectra after different incubation intervals in D_2O. These data enable the accurate examination of exchange rates of each assigned backbone amide. Importantly, the exchange rates provide additional information about the homogeneity and regularity within the fibrils (Luhrs et al., 2005, Ritter et al., 2005).

1.3.4. Solid-state NMR spectroscopy

Solid-state nuclear magnetic resonance spectroscopy (ssNMR) a powerful technique for structural investigations of insoluble proteins that do not easily form protein crystals or are not accessible to solution NMR spectroscopy, such as membrane proteins and amyloid fibrils. Several amyloid fibrils have been studied using ssNMR in recent years, isotope labelling of the fibrils is a requirement to obtain sequence-specific resonance assignments or to calculate distance torsion angles and restraints, and the high structural homogeneity is also extremely

important to get a good spectrum. The conformation of Alzheimer's Aβ fibrils has been investigated by ssNMR, the results showed that not only the intra-chain conformation of Aβ but also the inter-molecular alignment are explored the linkage of the mechanism of molecular association and the formation of Aβ fibrils (Luhrs, Ritter et al. 2005, Scheidt, Morgado et al. 2012). Atomic-resolution 3D structure of HET-s (218-289) fibrils examined by ssNMR (Ritter, Maddelein et al. 2005, Van Melckebeke, Wasmer et al. 2010, Van Melckebeke, Schanda et al. 2011), as well as other fibrils such as Ure2p (Baxa, Wickner et al. 2007) and Rnq1p (Wickner, Dyda et al. 2008) have been studied successfully.

2. Aims and Scope

Amyloids are discovered in bacteria, fungi, invertebrates, and humans and can play roles either in pathological or in physiological conditions. Amyloid proteins such as Aβ are associating with diseases. Almost all of disease-associated amyloid aggregates show cytotoxic. Other amyloids such as PMEL17 represent the native and functional states although these amyloids have similar biophysical features to disease-associated amyloids.

Culri is one of the well-studied extracellular functional amyloid fibrils produced in many enteric bacteria. Curli are essential for in the attachment, invasion of host cells as well as activation of the innate immune system. Curli is composed of a major subunit CsgA and a minor nucleator subunit CsgB. Recombinant CsgA fibrils represent a β-sheet seconday structure, however, there is no structural information at the atomic level for native Curli until now and the architecture of Curli is also poorly understood.

The aims of this study were the structural characterization of native Curli from *E.coli* and how CsgA and CsgB are assembled into artificial Curli *in vitro*. Firstly, In order to illustrate the structure of native Curli using solid-state NMR spectroscopy, isotope-labelled media were optimized and sufficient amounts of uniformly labelled Curli were produced. Secondly, mixed fibrils resulting from mixtures of CsgA and CsgB were studied using controlled shear-forces combined with seeding assays, and Immunogold labelling TEM as well as H/D exchange NMR technologies to reveal how CsgA and CsgB are assembled into mixed fibrils *in vitro*.

3. Material and Methods

3.1.Standard Materials

3.1.1. Chemical substances

Basic chemicals (summarized in **Table 4**) what were used in this work, all these substances are Research Grade and were purchased from Riedel, Sigma, PPA *et.al.*

Table 4: Chemical substances used in this work

Substances	Sources	CAS No.
Aceton	Riedel	67-64-1
Ammoniumchlorid	Riedel	12125-02-9
Bis-Tris	Sigma	6976-37-0
Bovine Serum Albumin	PAA	9048-46-8
Calciumchlorid	Merck	10043-52-4
Coomassie Brilliant Blue R250	Merck	6104-59-2
Congo Red	Roth	573-58-0
DL-Dithiothreitol	Sigma	3483-12-3
Dodecylsulfat Sodium Salt	Serva	151-21-3
EDTA	Fluka	6381-92-6
Ethanol	J.T.Baker	64-17-5
D(+)-Galactose	Merck	59-23-4
D(+)-Glucose	Merck	14431-43-7
Glycerol	Fluka	56-81-5
Glycine	Merck	56-40-6
Guanidine Hydochloride	Fluka	50-01-1
Imidazole	Fluka	288-32-4
IPTG	Carbosynth	367-93-1
Dipotassium hydrogen phosphate	Roth	7758-11-4
Magnesiumchlorid	Roth	7791-18-6
Sodium hydroxide	Merck	1910-73-2
Disodium phosphate	Roth	7782-85-6
Monosodium phosphate	Fluka	7758-16-9
Sodium azide	Merck	26628-22-8
Sodium chloride	Roth	7647-14-5
Sodium hydroxide	Merck	1310-73-2
Phosphoric acid	Sigma	7664-38-2
Protease Inhibitor Set	Roche	
Tricine	Sigma	5704-04-1
TCEP-HCl	Pierce	5961-85-3
Triton X 100	Sigma	9002-93-1
Tween 20	Biorad	9005-64-5
MEM Vitamins Solution(100x)	Sigma	

3.1.2. Oligonucleotide

Oligonucleotide (**Table 5**) was used to insert three amino acids-CGG to the N-terminus of CsgB to construct Cysteine modified CsgB $_{21\text{-}151}$ plasmids.

Table 5: Oligonucleotide used in this work

Primers	Sequence (5'-3')	Insert sites
CGG-CsgB	GAAGGAGATATACCATGTGCGGCGGAGC AGGTTATGATT	Insert CGG in N-terminal of CsgB $_{21\text{-}151}$

3.1.3. Recombinant plasmids

Recombinant plasmids were used, these plasmids (**Table 6**) derivative pET-11d amplification and expression in Escherichia coli.

Table 6: Plasmids used in this work

Plasmids	Features	Sources
pET-11d-CsgA $_{21\text{-}151}$	pET-11d derivative encoding *E. coli* CsgA $_{21\text{-}151}$	A. Eberth; HZI
pET-11d-CsgB $_{21\text{-}151}$	pET-11d derivative encoding *E. coli* CsgB $_{21\text{-}151}$	A. Eberth; HZI
pET-11d-N-Cys-CsgB $_{21\text{-}151}$	pET-11d derivative encoding *E. coli* N-Cys-CsgB $_{21\text{-}151}$	This work
pET-11d-N-Cys-CsgB $_{21\text{-}131}$ Δ132-151 CTRUNC	pET-11d derivative encoding *E. coli* N-Cys-CsgB $_{21\text{-}131}$ Δ132-151 CTRUNC	This work

3.1.4. Enzymes

High fidelity Pfu Turbo$^{\bigcirc;R}$ DNA polymerase was used to amplify gene templates. Restriction endonuclease DpnI was used to digest methylated plasmid afterwards (**Table 7**).

Table 7: Enzymes used in this work

Enzymes	Sources	Specification
DpnI	NEB	Restriction endonuclease
Pfu Turbo$^{\bigcirc;R}$Polymerase	Stratagene	DNA Polymerase

3.1.5. DNA and protein ladders

The DNA and protein ladders were used (**Table 8**).

Table 8: DNA ladder and protein ladders used in this work

Names	Sources
DNA Ladder Mix	Fermentas
PageRuler Prestained Protein ladder	Pierce
Unstained Protein ladder	Fermentas

3.1.6. Commercial Kits

The commercial Kits were used (**Table 9**).

Table 9: Commercial Kits used in this work

Kits	Sources	Usage
QIAquick©;® PCR Purification Kit	Qiagen	Purification of PCR products
QIAprep©;® Spin Miniprep Kit	Qiagen	Plasmids purification
BCIP/NBT kit	Promega	Detection of HRP conjugated Substrates
Novex©;®4-12 % Bis-Tris Gels	Invitrogen	Separation and detection of proteins

3.1.7. Bacterial strains

The bacterial strains were used for plasmids amplification, recombinant proteins expression and curli production respectively (**Table 10**).

Table 10: *E.coli* strains used in this work

Strains	Genotypes	Sources
E.coli DH5α	*F– endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG Φ80dlacZΔM15 Δ(lacZYA-argF)U169, hsdR17(rK- mK+), λ–*	Invitrogen
E.coli T7 Express	*fhuA2 lacZ::T7 gene1 [lon] ompT gal sulA11 R(mcr-73::miniTn10*-TetS*)2 [dcm] R(zgb-210::Tn10*--TetS*) endA1 Δ(mcrC-mrr)114::IS10*	NEB
E.coli MC4100	*F-, [araD139]B/r, Del(argF-lac)169, lambda-, e14-, flhD5301, Δ(fruK-yeiR)725(fruA25), relA1, rpsL150(strR), rbsR22, Del(fimB-fimE)632(::IS1), deoC1*	HZI

3.1.8. Buffers

The buffers and solutions were used in this study, are summarised in **Table 11**.

Table 11: Buffers used in this work

Buffers	Ingredients
Metal mixtures (1000 x)	100 µM $FeCl_3$, 40 µM $CaCl_2$, 20 µM $MnSO_4$, 20 µM $ZnSO_4$, 4 µM $CoCl_2$, 4 µM $CuCl_2$, 4 µM $NiCl_2$, 4 µM Na_2MoO_4, 4 µM H_3Bo_3
TAE buffer(1 x)	40 mM Tris, pH 8.2, 20 mM Sodium acetate, 1 mM EDTA
DNA loading buffer (6 x)	10 mM Tris-Cl pH 7.4, 25 mM EDTA, 30% (v/v) Glycerol, 0.4% (w/v) Orange G
Protein loading buffer(10x)	40% Glycerol, 240 mM Tris-Cl pH 6.8, 8% SDS, 0.04% Bromophenol blue, 5% β-mercaptoethanol
SDS-PAGE running buffer (1x)	25 mM Tris base, 192 mM Glycine, 0.1% SDS
12% SDS-PAGE gel(10ml)	4 ml Acrylamide/Bis-acrylamide (30%/0.8%w/v), 2.6 ml 1.5 M Tris-base pH 8.8,100 µl, 10% SDS, 100 µl 10% Ammonium persulfate, 10 µl TEMED, 3.2 ml ddH_2O
PBS(1x)	137 mM NaCl, 2.7 mM KCl, 8 mM Na_2HPO_4 , and 2 mM KH_2PO_4,pH 7.4
Transfer buffer	20 mM Tris base, 192 mM Glycine, 15% (v/v) Methanol, pH 8.0
Blocking buffer	20 mM Tris-Cl pH 8.0, 150 mM NaCl, 0.05% (v/v) Tween-20, 5% (w/v) Skimmed milk powder
Solubilization buffer	8 M Guandinine-Cl, 150 mM NaCl, 100 mM Kpi, 5 mM Imidazole, pH 7.4
Washing buffer	8 M Guandinine-HCl, 150 mM NaCl, 100 mM Kpi, 20 mM Imidazole, pH 7.4
Elution buffer	8 M Guandinine-HCl, 100 mM Kpi, 500 mM Imidazole, pH 7.4
Buffer A	50 mM Kpi, pH 7.4
Buffer B	5 mM Kpi, pH 3
Buffer C	20 mM Glycine,50 mM NaCl, pH 9.0

3.1.9. Media

The Mediums were used for Curli production and recombinant proteins expression respectively. These mediums are summarised in **Table 12**. All mediums were autoclaved at 121 °C for 30 min.

Table 12: Mediums used in this work

Mediums	Ingredients
LB*	1% Bacterial peptone, 0.5% Yeast extract, 1.2% Agar, pH 7.0
M9*	50 mM Na_2HPO_4 , 50 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose, Vitamin mixtures, Metal mixtures, 1.2% Agar, pH 7.0
CN-040	50 mM Na_2HPO_4, 50 mM KH_2PO_4, 1 g/ l NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 0.4% Glucose, Vitamin mixtures, Metal mixtures, 1.2% Agar, pH 7.0
N5052	50 mM Na_2HPO_4 , 50 mM KH_2PO_4, 1 g/ l NH_4Cl, 5 mM Na_2SO_4, 2 mM $MgSO_4$, 1.25% Glycerin, 0.125% Glucose, 0.5% Lactose, Vitamin mixtures, Metal mixtures, 1.2 % Agar, pH 7.0

N7050	50 mM Na_2HPO_4, 50 mM KH_2PO_4, 1 g/ l NH_4Cl, 5 mM Na_2SO_4 , 2 mM $MgSO_4$, 0.7% Glycerin, 0.05% Glucose, Vitamin mixtures, Metal mixtures, 1.2% Agar, pH 7.0
Modified M9	50 mM Na_2HPO_4 , 50 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
50 mM Na/K M9	50 mM Na_2HPO_4, 50 mM KH_2PO_4, 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
25 mM Na/K M9	25 mM Na_2HPO_4 , 25 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
10 mM Na/K M9	10 mM Na_2HPO_4, 10 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose , 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
1% Glucose M9	10 mM Na_2HPO_4, 10 mM KH_2PO_4, 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.1% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
2% Glucose M9	10 mM Na_2HPO_4, 10 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.2% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
4% Glucose M9	10 mM Na_2HPO_4 , 10 mM KH_2PO_4 , 1 g/ l NH_4Cl, 2 mM $MgSO_4$, 0.4% Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0
LB	1% Bacterial peptone, 0.5% Yeast extract, 1% NaCl, pH 7.4
^{15}N,^{13}C-modified M9*	10 mM Na_2HPO_4, 10 mM KH_2PO_4, 1 g/ l ^{15}N-NH_4Cl, 2 mM $MgSO_4$, 0.2% ^{13}C -Glucose, 100 µM $CaCl_2$, 1.2% Agar, pH 7.0

3.2. Methods

3.2.1. Expression and purification of ^{15}N,^{13}C-labelled Curli

Media optimization

The influence of environmental and physiological stress agents, e.g., salt concentration, nutrient limitation, temperature, desiccation and detergents had been studied to affect the yields of Curli in different strains of *E. coli*. In this study, MC4100 strain was used to produce native Curli. Various media were analyzed (**Table 9**). Na_2HPO_4/ KH_2PO_4 (50 mM, 25 mM and 10 mM) and glucose (0.1%, 0.2% and 0.4% (w/v)) concentrations were optimized. Incubation times were optimized as well.

Curli can show the red color when bound to amyloid specific dye-Congo-red. Curli formation of MC4100 bacterial cells was detected in various media (**Table 12**) in presence of 20 μM Congo-red, the redder of bacterial cells were indicated the higher yields of Curli.

Purification protocal

MC4100 bacterial cells were grown on ^{15}N, ^{13}C- labeled modified M9* agar plates at 28°C for 60 h before being scraped from 10 plates and suspended in 30 ml of cold PBS buffer. The suspension was processed for 10-15 min using the Water-bath (Elma Ultrasonic ○;R Cleaner Transsonic 310) to shear off Curli from the bacterial membrane. After the processed, the bacterial cells were pelleted by centrifuging at 4,000 g for 10 min at RT. The shorter fibrils in the supernatant were then collected by centrifuging at 16,000 g for 30 min at RT. The pellet (Pellet 1) was resuspended in PBS buffer containing 2% SDS and 1% Triton X-100, followed by gently shaking for 4-6 h at the room temperature before being centrifuged at 16,000 g for 30 min to get the pellet 2. The pellet (Pellet 2) was washed using PBS buffer for 3 times to remove residual SDS and Triton X-100. The last step was gel-purification; briefly, 0.25 ml of the fraction was mixed with an equal amount of 2×protein loading buffer and subjected to electrophoresis for 3-5 h on a 12% SDS-PAGE gel (The gel was parpared following by the standard SDS-PAGE gel protocol, and run as the standard protocol as well. The only difference was that the gel doesn't contain stacking gel). The Curli remaining in the slot after SDS-PAGE gel electrophoresis were recovered as described (Collinson, Emody et al. 1991). Curli sample（Pellet 3）was collected from the gel slot to wash 3 times using Milli-Q water. Unlabeled native Curli was also used the same method of purification (**Figure 7**).

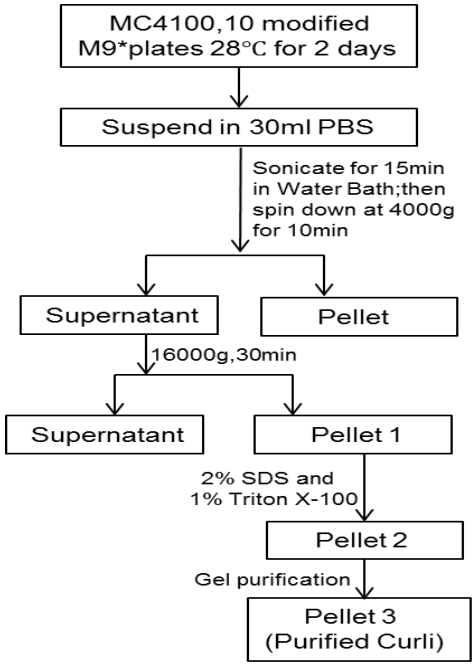


Figure 7: Purification procedure of native Curli isolated from MC4100 strain.

3.2.2. DNA and protein technologies

Multiple amino acids insertion

Oligonucleotide with TGCGGCGGA sequence encodes the amino acids of CGG was synthesized. Insertion of TGCGGCGGA sequence into N-terminal of $CsgB_{21\text{-}151}$.pET-11d plasmid was produced by PCR. The PCR condition is summarised in **Table 13**.

Table 13: Thermocycling conditions for site-insertion PCR

Segments	Cycles	Temperature	Time
1	1	95°C	2 min
		95°C	30 s
2	18	55°C	30 s
		68°C	2 min
3		68°C	5 min

The PCR product was treated by the restriction endonuclease DpnI at 37°C for 30 min resulting in digestion of the template plasmids, and then the digested PCR product was transformed into DH5α competent cells and incubated on a LB agar plate containing

ampicillin at 37°C for overnight. The second day, three single clones were picked out and incubated with 5-10 ml of LB medium with ampicillin overnight at 37°C respectively. The last step was isolated the plasmid, the plasmid was extracted using QIAprep$^{\bigcirc;R}$ Spin Miniprep Kit and sent to sequence. The modified CsgB plasmid with TGCGGCGGA sequence (**Figure 8**) was stored at -20°C for the further experiments.

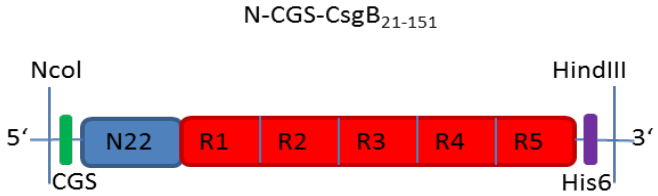


Figure 8: Insertion of TGCGGCGGA sequence into WT CsgB $_{21\text{-}151}$.

$CsgB_{21\text{-}151}$ constructs with flexible N-terminal region shown in blue with CsgB repeating units (R1-R5) are shown in Red. TGCGGCGGA sequence was inserted in the N-terminus shown in green.

Plasmids amplification

The plasmid was transformed into DH5α competent cells by heat shock at 42°C for 90 s, followed by that transformed DH5α Cells were incubated on LB agar plates (containing 100 µg/ml ampicillin) at 37°C for overnight. Afterwards, the single clone was incubated with 20 ml of LB medium (containing 100 µg/ml ampicillin) at 37°C for overnight shaking at 120 rpm. The QIAprep $^{\bigcirc;R}$ Spin Miniprep Kit was used for the plasmid extraction. NanoDrop ND-2000 can accurately read the concentration of DNA plasmids.

Expression and Purification of recombinant Proteins

Protein expression Competent T7 cells transformed with the plasmid (e.g CsgA $_{21\text{-}151}$-pET-11d) was grown on a LB plate (100 µg/ml ampicillin) for overnight at 37°C. The next day, a single colonly was incubated into 20 ml LB medium (100 µg/ml ampicillin) and the pro-culture was grown to an OD_{600} of 1.8-2 at 37°C with shaking. The pro-culture was mixed with 1L LB medium (100 µg/ml ampicillin), and the culture was further grown to 0.8-0.9 of OD_{600} at 37°C with shaking. Isopropyl-1-thio-B-D galactopyanoside (IPTG) was added to the final concentration of 0.5 mM, and the culture was further incubated for 4-5 h at 37°C. After induction, the bacterial cells were harvested by centrifugation for 10 min at 4,000 g. And the

bacteria were stored at -80°C for the further experiments.

1 L culture of the bacteria cells were suspended in 250 ml Solubilization buffer and incubated for 18-24 h with stirring at room temperature, and then centrifuged the suspension at 16,000 g for 30 min at room temperature. The supernatant was collected and then incubated with the 2-3 ml Ni-sepharose beads equilibrated with Solubilization buffer for 1 h at room temperature with slight stirring. The Ni-sepharose beads bound with recombinant his-tagged proteins were collected by 2,000 g for 5-10 min. 100-150 ml of washing buffer was used to wash the Ni-sepharose beads to remove the contaminants efficiently. The last purification step was that recombinant his-tagged protein was stripped from Ni-sepharose beads with 4-6 ml Elution buffer. Recombinant CsgA or CsgB proteins were very stable stored in elution buffer for 5-7 days at room temperature or stored at -80°C for more than 2 months.

SDS-polyacrylamide Electrophoresis

Discontinuous precast 4-12% SDS-polyacrylamide (SDS-PAGE) gel was used to separate proteins to determine proteins molecular weight and purity. Running of SDS-PAGE was performed according to NuPAGE$^{\bigcirc;R}$ SDS-PAGE gel protocol. For proteins preparation, soluble protein in specific buffer was directly mixed with protein loading buffer. However, insoluble fibril sample must be dissolved in 98% formic acid firstly and dried it using vacuum freeze-drying. The treated protein sample was further mixed with protein loading buffer for SDS-PAGE experiment.

The mixture of protein sample and loading buffer were boiled at 95°C for 10 min and immediately was loaded on 10 µl samples per lane as well as 2 µl protein ladders, gel-running condition was at 200 V of constant voltage for 40 min. Finally, SDS-PAGE gel was stained with approximately 10 ml Instant Blue^{TM} (Expedeon) for 5-10 min incubation at room temperature with gentle shaking.

Western blot

For protein immunoblot, Protein samples were firstly separated based on molecular weight using SDS-PAGE gel electrophoresis. These proteins were then transferred to and immobilized on a Nitrocellulose or Polyvinylidene fluoride (PVDF, Millipore) membrane. The transferred PVDF membrane was incubated with specific antibody of the protein of interest. Briefly, the filter paper (Whatman) was wetted in cold transfer buffer, and

PVDF membrane was activated in cold methanol at the same time. Followed to create a transfer "sandwich" (Filter paper- gel-PVDF –Filter paper), and then transferred the proteins from SDS-PAGE gel to PVDF membrane using a semi-dry blot apparatus, transferring was carried out at 18 V for appropriately 20-30 min depending on the size of proteins. After semi-dry blot, the transferred membrane was washed using TBS buffer for 3 times and then blocked the membrane with 5% Skim milk in TBS for 1 hour at the room temperature, the membrane was incubated with primary antibody (1:2000 dilution) for overnight at 4°C with gentle shaking. The second day the membrane was washed for 3 times using TBS/T buffer and then incubated with secondary antibody (1:5000 dilution) for 60 min at room temperature with gentle shaking. The last step was that the immobilised proteins were detected by BCIP/NBT for 3 min at room temperature.

3.2.3. Buffer exchange abnd Fibrilisation

Buffer exchange

All recombinant his-$CsgA_{21-151}$, his-$CsgB_{21-151}$ and Cysteine-CsgB protein were purified and stored in Elution buffers with their denaturing states. These proteins were buffer-exchanged into pH 7.4 buffer (Buffer A) ($CsgA_{21-151}$) or pH 7.4 buffer (Buffer B) ($CsgB_{21-151}$ and Cysteine-CsgB) to produce the fibrils.

The protein sample was firstly filtered by 50kD of Vivaspin (Sartorius Stedim) to remove any protein aggregates, and then the sample was loaded on an equilibrated NAP-5 or PD-10 desalting column (GE Healthcare) to buffer-exchange the protein into pH 7.4 buffer or pH 7.4 buffer (**Table 14**).

Table 14: Buffer exchange using NAP-5 and PD-10 desalting column

Desalting Column	Equilibration	Sample volume	Elution volume
NAP-5	10-20 ml buffer	500 µl	1 ml buffer
PD-10	50-100 ml buffer	2.5 ml	3.5 ml buffer

Fibrillation of CsgA protein

Filtered CsgA protein was buffer-exchanged into pH 7.4 buffer (As described in 3.2.3) and

then the CsgA was diluted to final concentration of 20 μM with 0.05% sodium azide. For CsgA fibrillation, the CsgA was incubated for uncontinuous shaking (shaking 10 s every 10 min intervals, 900 rpm) by Thermomixer (Eppendorf) at the RT for at least 2 days.

Fibrillation of CsgB protein

Filtered CsgB protein was firstly buffer-exchanged into pH 3 buffer (As described in 3.2.3) because of avoiding premature aggregation, and then CsgB was immediately diluted to final concentration of 20 μM with 0.05% sodium azide after being adjusted the buffer pH from pH 3.0 to pH 7.4 (using 1 M H_3PO_4). For CsgB fibrillation, the CsgB was incubated with uncontinuous shaking (shaking 2 s every 10 min intervals, 900 rpm) by Thermomixer (Eppendorf) at the room temperature for 2 days.

Fibrillation of mixtures CsgA protein and CsgB protein

Filtered CsgA protein and CsgB protein were buffer-exchanged into pH 7.4 buffer and pH 3 buffer respectively (As described in 3.2.3). For fibrillation, CsgA protein and CsgB protein were immediately mixed with the specific concentrations in presence of 0.05% sodium azide and adjusted buffer pH to 7.4 (using H_3PO_4). The protein mixtures were incubated to form mixed fibrils at room temperature for 2 days.

3.2.4. Physical methods to study the fibrils

Thioflavin T fluorescence assay for CsgA aggregation

The aggregation kinetics of CsgA protein was measured by monitoring the change of fluorescence intensity of Thioflavin T (ThT) upon dye binding monitored using a Plate Reader (TecanM1000). Filtered CsgA protein was desalted into pH 9 buffer (or pH 7.4 buffer). The reaction mixture was composed of 10 μM Csg A mixed with or without specific seeds in presence of 20 μM ThT dye.

The mixtures (200 μl for each well) were loaded onto a 96-well flat bottom black polystyrene plate (Corning 96 Flat Bottom Black Polystyrol) and the plate was carefully covered using Sealing Tape (Nalge Nunc). The plate was then immediately transferred into a Plate-Reader. ThT dye fluorescence was measured and set to excitation/emission at 442/485 nm. Measurements were taken by top reads every 10 min, and shaking for 5 s at 400 Hz at 25°C. Reads were integrated for 20 μs, a gain of 105 and multiple reads per well (5*5).

Circular Dichroism

Protein Secondary structure can be determined in the far-UV spectral region from 260 nm to 190 nm. At these wavelengths the chromophore is the peptide bond, α helix, β sheet and random coil structures each give rise to a characteristic shape and magnitude of the CD spectrum. For far-UV CD spectra, soluble protein was required to adjust to 1 mM peptide bonds with CD suitable buffer that does not have absorbance in far UV region. Amyloid fibrils were sonicated briefly to minimize fibril precipitation before measurement. The sample was applied to cuvettes with a path length of 1 mm. Spectra was taken using a J-815 CD spectrometer (JASCO) and all spectra were baseline corrected with respect to buffer. To improve the signal-to-noise ratio, 3 scans were averaged for each sample. Measurement was performed in continuous scanning mode with a resolution 0.2 nm at 10°C.

Fourier Transform Infrared Spectroscopy

Fourier Transform Infrared Spectroscopy (FT-IR) spectroscopy is a measurement of wavelength and intensity of the absorption of Infrared spectroscopy radiation by protein sample. The most sensitive spectral region to the protein secondary structural components is named- the amide I region (1690-1600 cm^{-1}). Secondary structure of α helix, β sheet and random coli give rise to a characteristic C=O stretching frequencies due to unique molecular geometry and hydrogen bonding pattern. FTIR spectroscopy was conducted using a Tensor 27 (Bruker) equipped with a Bio-ART cell. The fibril samples were washed at least 3 times with Milli-Q water and adjusted the sample concentration to 1 mg/ml in specific buffer. Spectra were recorded from 1900-1500 cm^{-1} using mid-IR source and KBr beam splitter, 100 scans were averaged for each sample. All data were processed by OPUS 5.5 software.

Mass Spectrometry

Mass spectrometry (MS) is an important method for the characterization of proteins. Matrix assisted laser desorption ionization (MALDI), a main method for ionization of proteins, is being applied to analyze protein molecular weight. Soluble protein samples were directly desalted into deionized water and adjusted to the final concentration of 0.2-1 mg/ml. The fibrils samples (20-50 μg) were dissolved in 50 μl of 98% (v/v) formic acid. In preparation of MALDI-MS samples, 0.5 μl of protein solution (1-10 pmol/μl) was loaded onto a MALDI

plate and then immediately mixed with equal µl of 30 mM 2, 5-Dihydroxybenzoic acid (DHB) or α-cyano-4-hydroxycinnamic acid (α-cyano) matrix solution and the mixture were set to dry at room temperature. Afterwards, the dried sample/matrix was washed with cold Milli-Q water for removing salts. The molecular weight of samples was measured using Ultraflex Time-of Flight mass spectrometer (Bruker) of the positive mode. Ms. Anja Meier (HZI) performed all MALDI-MS measurements.

Transmission Electron Microscopy

Transmission Electron microscopy (TEM) was used to probe the morphology of amyloid fibrils.

(1) Biotin-labelled mixed CsgA-CsgB fibrils

Before N-CGS-$CsgB_{21\text{-}151}$ was treated with the EZ-Link Maleimide PEG2-Biotin (Thermo) reagent, potential disulfide bonds were reduced by incubation with 5 mM TCEP for at least an hour at room temperature. The protein was then buffer-exchanged into 8 M Guandinine-HCl, 100 mM Kpi, pH 7.4 buffer to remove the excess TCEP using a NAP-5 column. According to protein concentration, a 5-fold molar excess of EZ-Link Maleimide PEG2-Biotin was added to the protein sample and kept on the rotating wheel for 5 h or overnight at room temperature. After biotin labeling, the N-CGS-$CsgB_{21\text{-}151}$ CsgB protein was buffer-exchanged into pH 3 buffer. In order to check the efficiency of the labelled reaction, the molecular weight was measured immediately using MALDI-MS.

Mixed fibrils of CsgA and biotin-labelled N-CGS-CsgB were immediately afterwards produced as described in **3.2.3**. Afterwards, 100 µl of the mixed CsgA-CsgB fibrils were incubated with 2-4 µl of Gold labelled anti-biotin antibody (Size of gold particle: 10 nm**)** for an hour at room temperature, followed by washing with pH 7.4 buffer or ddH_2O to remove the excess Gold-labelled anti-biotin antibody.

(2) Negative-stained TEM

Prof. Dr. Manfred Rohde (HZI) performed the uranly acetate staining of fibrils.

TEM was performed using Zeiss 910 transmission electron microscope (Zeiss) instrument operating at 80 kV. Images were taken using slow scan CCD camera（Proscan）with Olympus ITEM image software. Prof. Dr. Manfred Rohde (HZI) took all EM pictures.

Solid-state NMR analysis of native Curli

For the solid-state sample preparation, uniformly ^{15}N,^{13}C-labelled native Curli was washed three times using Milli-Q water, and the Curli sample was further centrifuged (14,000 rpm for 30 min) into the 3.2 mm rotor (Bruker) using a specially designed filling device (Gardiennet, Schutz et al. 2012). Approximately 25 mg of uniformly labeled Curli was filled.
A solid-state ^{13}C-^{13}C NMR measurement of uniformly labeled native Curli was performed at an Avance 900 NMR spectrometer (Bruker) in FMP-Berlin, equipped with four radio-frequency channels and a 4-mm double-resonance MAS probe. The two-dimensional ^{13}C-^{13}C correlation spectrum was acquired at MAS frequencies of 13.5 kHz and measured with the ^{13}C-^{13}C mixing accomplished using though-space dipolar assisted rotational resonance (DARR). The mixing time is 25 ms and Acquisituon times are 15.36 ms (F2) and 8.448 ms (F1). The spectrum was processed using TopSpin 3.2 (Bruker) and analyzed using CcpNmr analysis (version 2.4).

H/D exchange NMR analysis of mixed CsgA-CsgB fibrils

Mixed CsgA-^{15N}CsgB (80%:20%) fibrils were studied by H/D exchange NMR experiment.

(1) **Preparation of mixed fibrils**

CsgA and ^{15}N-CsgB were buffer-exchanged into pH 7.4 buffer and pH 3 buffer respectively. Mixed fibrils were fibrillized at the final concentration of 20 μM and were composed of 80% CsgA and 20% ^{15}N-CsgB for 2 days as described in **3.2.3**-Fibrillation of mixtures CsgA protein and CsgB protein. Mixed fibrils were then incubated with 1-2% SDS or 6 M urea overnight at room temperature to wash away any unfolded or oligomers proteins (Wang, Smith et al. 2007), followed by washing 3 times with pH 7.4 buffer to remove residual urea or SDS. This procedure enhanced the population of specific ordered mixed fibrils.

(2) Preparation of H/D exchange samples

Aliquots of 30mg mixed fibrils (wet weight) were washed once with 200 μl of D_2O buffer (50 mM Kpi in D_2O, pH 7.4), centrifuged at 16,000 rpm for 3 min and the D_2O buffer removed by a pipet carefully . Subsequently, the pellets were incubated with 350 μl of D_2O buffer in order to facilitate H/D exchange between the mixed fibrils and D_2O. The H/D exchange reactions were quenched at specific time points (0 min, 30 min, 1.5 h, 8 h, 24 h, 1w, 2 w, 4 w and 6 w) by removing D_2O buffer and freezing in liquid nitrogen.

(3) Spectra recording and data analysis

The H/D exchanged mixed fibrils was immediately monomerised and dissolved in 550 µl of a mixture of DMSO-d6 and 0.1 % TFA-d1, meanwhile, the NMR setting was pre-adjusted. The total time of the sample preparation and the NMR setting must be controlled less than 4 min (3 min 50 s was used). The series of up to 5 2D- [^{15}N, ^{1}H] correlation spectra (each spectra corresponding to 10 min) were immediately recorded.

For H/D exchange data analysis, the spectra were processed using Topspin 3.2 (Bruker). And then, all 5 of processed spectra for each quenched time points were summed together for intergration. In order to determine the specific exchange rates, the cross-peak intensities (height type) of each residues in CsgB were calculated using CcpNMR analysis (version 2.4), these intersities were further calculated using Metlab with the help of Dr. Gang Zhao (HZI), and fitted either monoexponential or double exponential to the exchange time.

A single exponential decay function (equation 1) and a double exponential decay function (equation 2) were fitted the the time-dependent, relative cross peak intensities,

$$Y(t) = a \cdot \exp(-k_{ex} \cdot t) + c \quad \text{(Equation1)}$$

$$Y(t) = a \cdot \exp(-k_{ex1} \cdot t) + b \cdot \exp(-k_{ex2} \cdot t) + c \quad \text{(Equation2)}$$

Exponential decay: a and b are the pre-exponentual factors, c is the peck intersity at infinite time, k_{ex}, k_{ex1} (fast exchange) and k_{ex2} (slow exchange) are the amine proton exchange rate, t is the time between initiation of H/D exchange and quenching of this exchange.

To identify the relative population of residue showing biphasic behavior,the below equation was used. a and b are the pre-exponentual factors.

$$\text{Relative population (major population)} = (a-b) / a$$

Analysis of amyloid fibrils by Selective Shearing Amplification array

Selective Shearing Amplication array (invented by Dr. Thorsten Lührs) is kind of instrument that has been used to produce aggregated state prion protein from native-conformational prion protein. The principle of Selective Shearing Amplication array is based on the use of mechanical shear force as a shearing mechanism. Briefly, when the viscous fluid is confined

in the space between two rotating cylinders, homogenous shear fields can be produced that are proportional to the rotation frequency of the cylinder (Vandana Gupta, PhD thesis). “long” fibrils are in this homogenous shear fields that can be fragmented into “short” fbrils **(Figure 9)**. In this study, Selective Shearing Amplication array was used to study the characteristics of mixed CsgA-CsgB fibrils. For fibrils preparation, the fibrils have to be washed with 1% SDS twice for removing the non-specific aggregates and followed by washing with pH 7.4 buffer three times to remove the residual SDS. The reaction concentration of the fibrils was 5 µM in pH 7.4 buffer using a volume of 1 ml per reaction (**Table 15**).

Before the process, the rotors have to completely clean with the reaction buffer. The fibrils were subjected to various cycles (3, 6 and 12 cycles) of 5 s rotation and 30 s resting phase without agitation at 37°C. The rotation frequencies were controlled from 240HZ to 960HZ (240HZ, 300HZ, 380HZ, 480HZ, 600HZ, 760HZ and 960HZ) with an accuracy of rotation of 0.5-5HZ. The processed fibrils were further used to seed freshly monomer CsgA protein immediately.

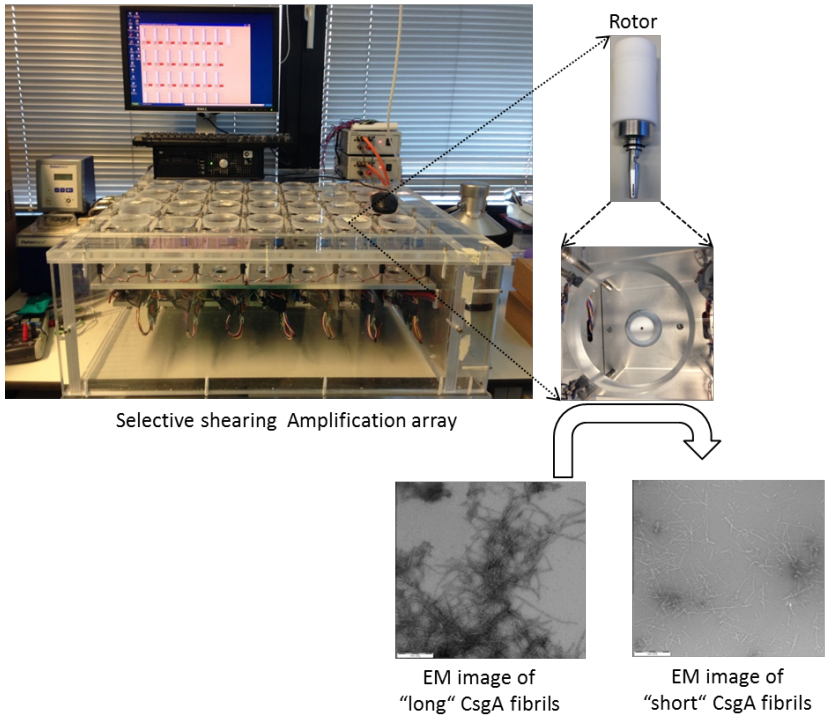


Figure 9: Selective Shearing Amplification array

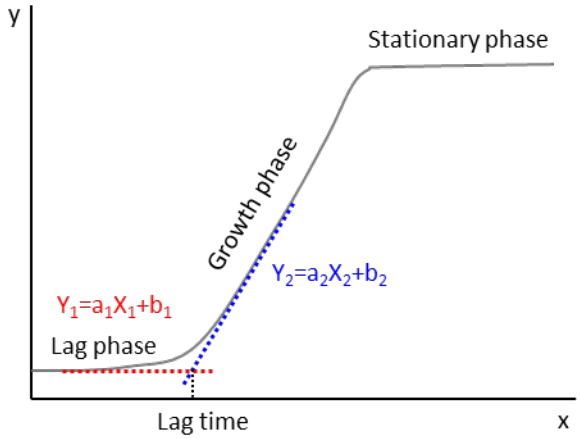
Selective Shearing Amplification array is the homemade apparatus; the rotation of rotor is controlled by the software. “long” CsgA fibril can be fragmented into “short” CsgA fibrils at specific condition (1200 Hz, 3 min). The fibrils were examined by TEM, scale bar is 200 nm.

Table 15: Fibrils were processed by Selective Shearing Amplification array

Fibrils	Reaction concentration	Reaction volume
CsgA fibrils	5 μM	1 ml
Mixed CsgA-CsgB (95%:5%) fibrils	5 μM	1 ml
Mixed CsgA-CsgB (90%:10%) fibrils	5 μM	1 ml
Mixed CsgA-CsgB (80%: 20%) fibrils	5 μM	1 ml

Caculation of the lag time

The curve of CsgA aggregation can be divided into three phases, lag phase, growth phase and stationary phase (**1.2.6**).



In order to calculate the lag time, firstly, to darw a CsgA aggregation's curve using Excel, the x axis represents the time and the y axis represents the ThT fluorescence intensity.

Secondly, to select several time points (10-15) in lag phase and in growth phase respectively (See above figure), and further fit these selected time points to a linear equation $Y_1=a_1X_1+b_1$ (lag phase) and $Y_2=a_2X_2+b_2$ respectively, wherein a_1, b_1, a_2 and b_2 are constants.

Thirdly, when $Y_1 = Y_2$ and $X_1 = X_2$, the below equation can be used to calculate the lag time (X represents the lag time in the equation).

$$X= (b_2 - b_1) / (a_1 – a_2)$$

4. Native Curli displays a well-ordered β-sheet core

Curli and cellulose are two of the most important components of the extracellular matrix biofilms of *E. coli*, which promote bacterial attachment to host cell surfaces (Saldana, Xicohtencatl-Cortes et al. 2009, Macarisin, Patel et al. 2012). The *E.Coli* strain MC4100 bacterial cells can produce a good yield of Curli but produces little or no cellulose; therefore, MC4100 can be used as the suitable strain to produce native Curli (Chapman, Robinson et al. 2002). Furthermore, Curli yield is tightly regulated by many environmental factors such as growth temperature, salts and nutrients of media (Romling, Bian et al. 1998, Jubelin, Vianney et al. 2005).

Production of amyloid fibrils for structural studies usually involves the *in-vitro* fibril formation, which can lead to heterogeneity or polymorphism. This was observed for the disease associated Aβ and α-synuclein (Townsend, Shankar et al. 2006, Auluck, Caraveo et al. 2010, Li, Dolios et al. 2014). In an earlier solid-state NMR study concluded that *in-vitro* fibrilized CsgA is not based on an in register β-sheet (Shewmaker, McGlinchey et al. 2009). In comparison to *in-vitro* fibrilized CsgA, Curli are the natural fibrils produced by *E.coli*, and show therefore exhibit the physiological structural features. In order to illustrate the structural information of native Curli using solid-state NMR spectroscopy, sufficient amounts of isotope-labelled Curli have to be produced. The aims of this part of the thesis were therefore:

(1) Optimization of the medium for uniformly isotope labeled native Curli production;

(2) Establishment of an optimized method for the purification of native Curli based on published protocols;

(3) Investigation of the structural features of native Curli compared to *in-vitro* fibrilized CsgA fibrils using solid-state NMR spectroscopy.

4.1. Media optimatization

MC4100 bacterial cells were previously used to produce native Curli grown on LB plates without salt (termed Modified LB) at 28°C for 2 days (Chapman, Robinson et al. 2002). However, Modified LB can't be used to produce the uniformly labeled Curli becease its ingredients cannot be isotope labelled. In this study, MC4100 bacterial cells were used to

produce native Curli, and the Curli yields were firstly detected on five different isotope-labeled media (M9, Modified M9, CN-040, N5052 and N7050), containing 40 μg/ml of Congo red dye (**Table 12**). Congo red can bind to amyloid, giving rise to red color in bacterial colonies producing extracellular amyloid. In other words, the binding capacity of Congo red to bacterial cells reflects the level of Curli production (Chapman, Robinson et al. 2002, Nenninger, Robinson et al. 2011, Hammer, McGuffie et al. 2012).

As shown in **Figure 10**, MC4100 bacterial cells showed a white color on the Congo red indicator plates of CN-040, N5052 and N7050 respectively, and bacterial cells exhibited a pale red color grown on both M9 and Modified M9 plates. In contrast, bacterial cells showed dark shade of red when grown on the control plate of Modified LB. The results indicated that bacterial cells produce no or very little Curli when grown on CN-040, N5052 or N7050 and lower amount of Curli when grown on M9 or Modified M9 compared the control medium-Modified LB. Thus, Modified M9 medium was used for further experiments.

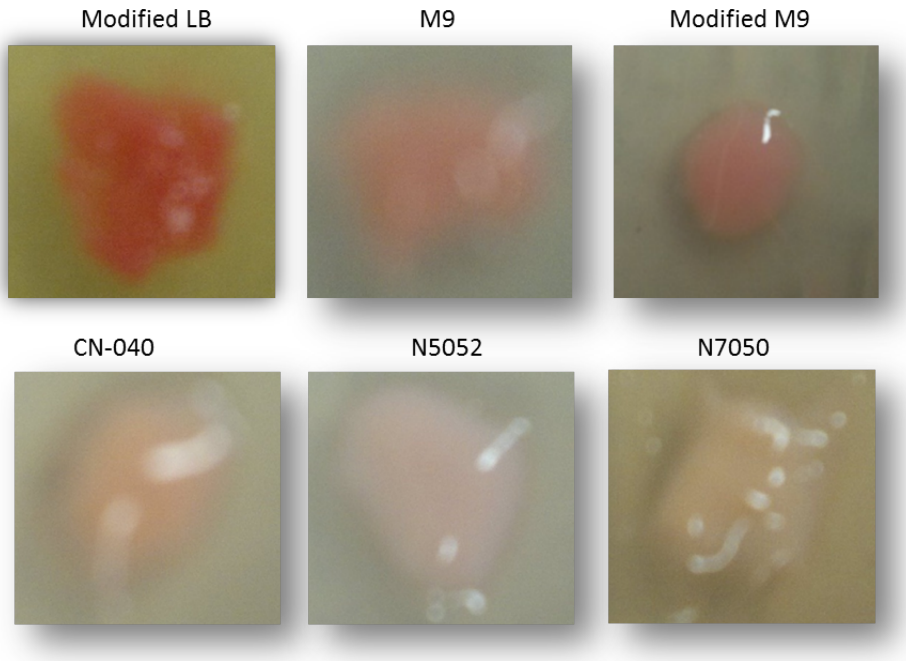


Figure 10: Production levels of native Curli for 6 different growth media

Phenotypic presentation of *E.coli* MC4100 bacterial cells grown on 6 media (Modified LB, M9, Modified M9, CN-040, N5052 and N7050) supplemented with Congo red (40 μg/ml) at 28°C for 2 days. The intensities of colony pigmentation as a result of Congo red binding indicate different levels of Curli production.

When *E.coli* is grown on media with the high concentration of salts, the CpxR regulator represses the expression of csgD resulting in a decreased Curli production (Jubelin, Vianney et al. 2005). In order to improve the yield of Curli, the salt ingredients of Modified M9

medium were modified. In detail, the Na_2HPO_4/KH_2PO_4 concentration was reduced to 10 mM (10mM Na/K M9 medium) and 25 mM (25 mM Na/K M9 medium) respectively (**Table 12**). As shown in **Figure 11**, MC4100 bacterial cells were grown on 10 mM Na/K M9, 25 mM Na/K M9 and 50 mM Na/K M9 plates respectively, each supplemented with 40 µg/ml of Congo red. They were incubated for 1 day, 2 days, 3 days and 4 days respectively at 28°C.

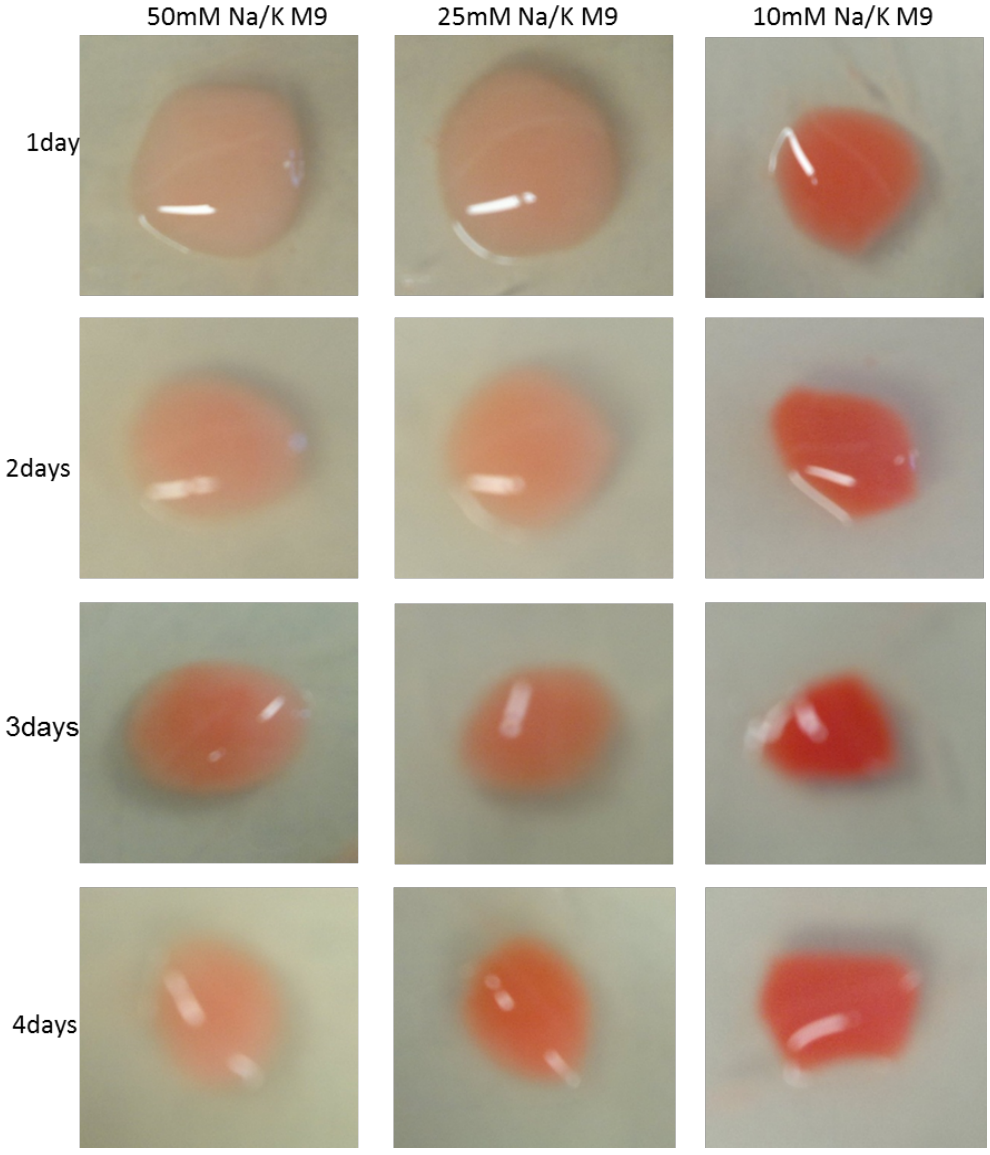


Figure 11: Production levels of native Curli dependent on incubation times and salt concentrations

Phenotypic presentation of *E.coli* MC4100 strain grown on 50 mM Na/K M9, 25 mM Na/K M9 and 10 mM Na/K M9 media respectively supplemented with Congo red (40 µg/ml) at 28°C for 1 day, 2 days, 3 days and 4 days respectively. The intensities of colony pigmentation as a result of Congo red binding indicate different level of Curli production.

On the first day, in comparison to bacterial cells grown on 25 mM Na/K M9 and 50 mM

Na/K M9 plates, bacterial cells displayed the darkest shade of red in 10 mM Na/K M9 plate. These results indicated that bacteria produced higher amounts of Curli grown on low concentrations of phosphate salts. On the second day, bacterial cells on all the plates showed the quite similar colors compared with their colors of the day before. On the third and fourth day, bacterial cell on all the plates were getting redder, and bacterial cell on the 10 mM Na/K M9 plate displayed the darkest shade of red compared with the two other plates. Sum up briefly, bacterial cells produced the highest yield of Curli grown on 10 mM Na/K M9 plates for 2-3 days, this result is because of the lowest concentrations of Na_2HPO_4 and KH_2PO_4.

Except for salts, glucose is another important component that could affect the yield of native Curli. Thus, the glucose concentration in 10mM Na/K M9 medium was modified to get three new media of 0.1% Glucose M9 medium, 0.2% Glucose M9 medium and 0.4% Glucose M9 medium (**Table 12**). As shown in **Figure 12**, these media at 28°C for 2 days. In comparison to bacterial cells grown on 0.1% Glucose M9 and 0.4% Glucose M9 plates, bacterial cells displayed the darkest shade of red grown on 0.2% Glucose M9. These results indicated that bacterial cells produced the highest yield of Curli grown on 0.2% Glucose M9 plate than the two other plates.

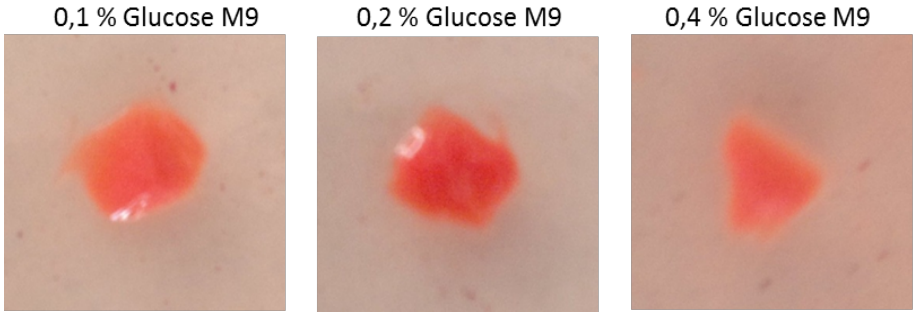


Figure 12: Production levels of native Curli dependent on glucose concentrations

Phenotypic presentation of *E.coli* MC4100 cells grown on 0.1% Glucose M9, 0.2% Glucose M9 and 0.4% Glucose M9 media respectively supplemented with Congo red (40 μg/ml) at 28°C for 2 days. The intensities of colony pigmentation as a result of Congo red binding indicate different levels of Curli production.

The optimized conditions for isotope labelled Curli production are following:

Medium(^{15}N,^{13}C-modified M9*)	Incubation temperature	Incubation time
10 mM Na_2HPO_4, 10 mM KH_2PO_4, 1 g/ l ^{15}N-NH_4Cl, 2 mM $MgSO_4$, 0.2% ^{13}C -Glucose, 100 μM $CaCl_2$, 1.2% Agar, pH 7.0	28°C	2 days

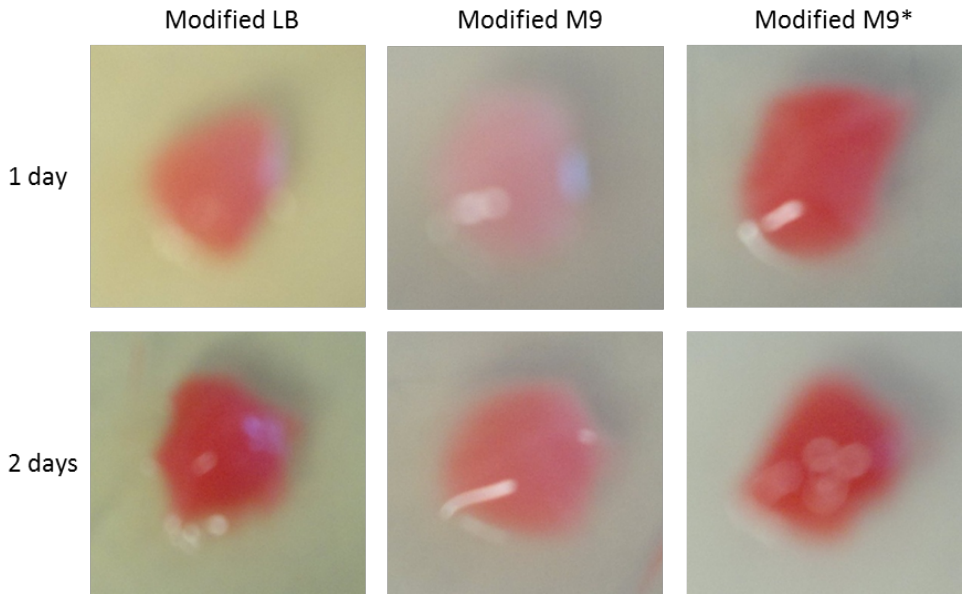


Figure 13: MC4100 bacterial cells grown on Modified LB, Modified M9 and Modified M9* respectively

Phenotypic presentation of *E.coli* MC4100 bacterial cells grown on Modified LB, Modified M9 and Modified M9* media supplemented with Congo red (40 μg/ml) at 28°C for 1 day and 2 days respectively. The intensities of colony pigmentation as a result of Congo red binding indicate different level of Curli production.

In conclusion, MC4100 bacterial cells produced the highest yield of native Curli grown on the optimized medium of Modified M9* (also termed 0.2% Glucose M9) compared with the medium of Modified M9, and the yield reached or excessed the Curli yield of Modified LB-the initial control medium (**Figure 13**). These results demonstrated that MC4100 bacterial cells can produce much more native Curli grown on the medium with the lower salts and glucose.

4.2. Purification of isotope-labeled native Curli

The native Curli purification method used in the thesis is based on published protocol (Collinson, Emody et al. 1991). The most important diffence is that Curli are sheared off by the water-bath. The procedure of Curli purification is shown in **Figure 14A.** Briefly, The MC4100 bacterial cells were grown on ^{15}N,^{13}C-labelled modified M9* plates. Bacterial cells were afterwards scraped from the plates and suspended in cold PBS buffer, then bacterial cells were processed in the water bath, in this step, only Curli fibrils and less outer membrane proteins were sheared off the bacterial surface while leaving the bacterial cells intact because of low intensity of the ultrasound in the water-bath.

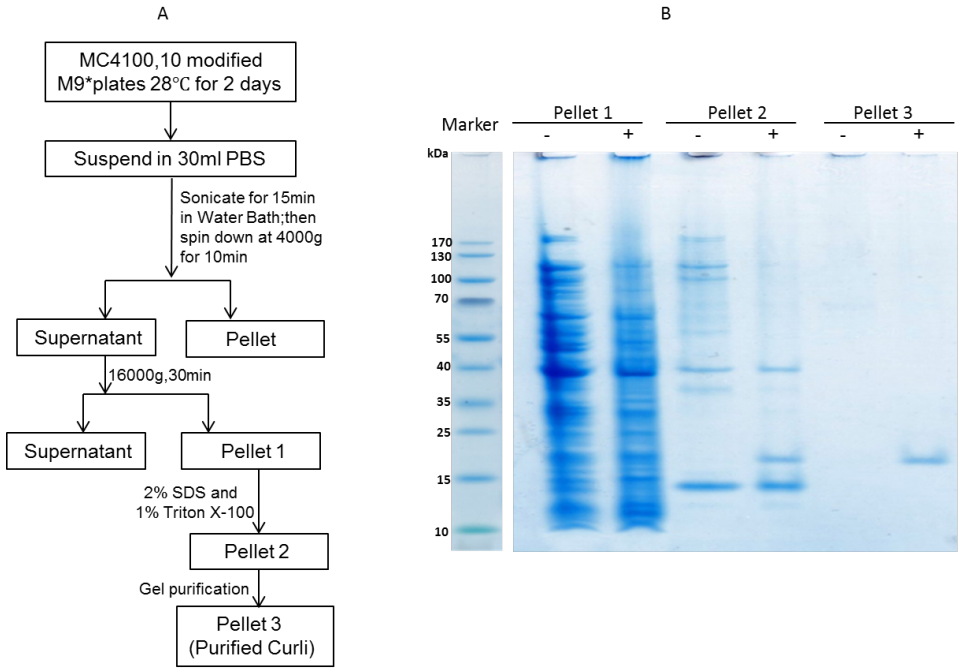


Figure 14: Purification of isotope labelled native Curli from MC4100 bacterial cells

A: The purification procedure of isotope-labelled native Curli. Three intermediates (Pellet 1, Pellet 2 and Pellet 3) were analyzed in order to evaluate native Curli purity. Pellet 1 (Crude sample was treated with sonication); Pellet 2 (the sample was treated with 2% SDS and 1% Triton X-100) and Pellet 3 (the sample was treated with Gel-purification).
B: These crude intermediates were treated with (+) or without (-) 98% formic acid respectively, and then analyzed by SDS-PAGE gel. The molecular weight of native Curli should be around 14 kDa. However, the native Curli bands of all of these samples run at approximately 18 kDa. Pre-stained protein ladder is indicated on the left.

After sonication, the bacterial cells were pelleted by centrifugation. The supernatant was carefully collected, and from this a crude Curli preparation (Pellet 1) was pelleted by centrifugation. Pellet 1 contained Curli and contaminants (**Figure 14B**). In order to remove the contaminants, Pellet 1 was treated with SDS and Triton X-100, because contaminants contain a larger amount of outer membrane proteins that can be desolved in SDS or Triton X-100.

Pellet 2 was collected by centrifugation (**Figure 14A**). Because untreated Curli cannot enter the gel, Curli had to treat with 98% formic acid (FA) to depolymerize into the monomeric state of CsgA. The CsgA monomer presents a band of around 18 kDa. SDS-PAGE analysis showed that the Curli sample of Pellet 2 was more pure than Pellet 1(**Figure 14B**).

To further enhance the purity of Curli sample, Pellet 2 was further purified using Gel-purification (Collinson, Emody et al. 1991). Because Curli don't dissolve in loading buffer and cannot enter the gel, the Curli remained in the slot after electrophoresis; however, almost all contaminants enter the gel. Pellet 3 was recovered from the slot. SDS-PAGE gel result

showed untreated Pellet 3 presented nothing; however, treated Pellet 3 presented only one band of around 18 kDa (**Figure 14B**).

4.3. Biochemical properties of native Curli

Native Curli was isolated from MC4100 bacterial cells as described in the previous section. The purified Curli was next examined by Western blot, negative-stain EM and FT-IR.

In order to determine the subunit composition of the purified Curli, the purified fibrils were analyzed by western blot using anti-mouse CsgA and anti-mouse CsgB antibodies, the result showed that native Curli contained approximately 90-95% CsgA and 5-10% CsgB (**Figure 15A**).

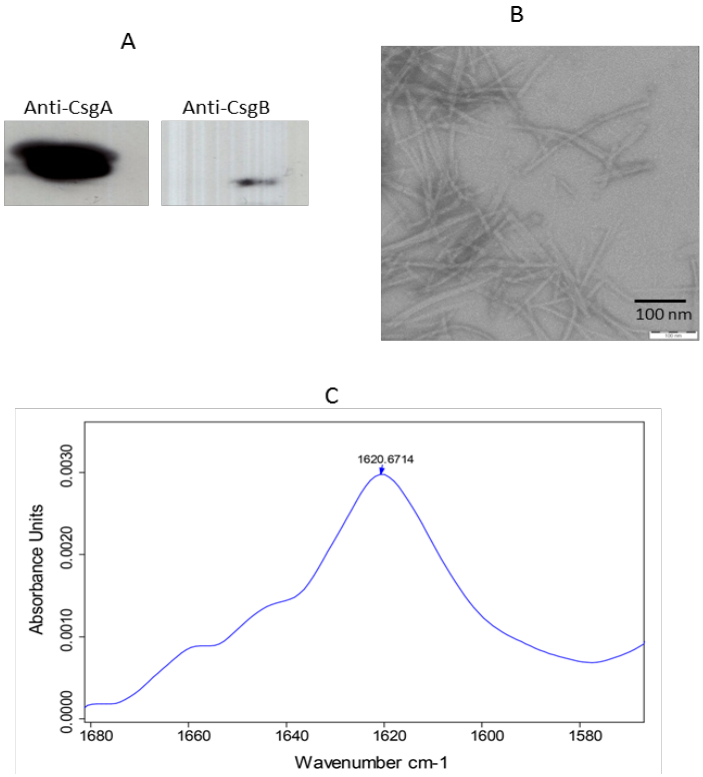


Figure 15: Biochemical properties of native Curli isolated from MC4100 bacterial cells

Native Curli isolated from MC4100 bacterial cells were grown on modified M9* plates at 28°C for 2 days. A: Western blot for purified Curli, both anti-mouse CsgA and CsgB antibodies against purified Curli. B: Negative-stain EM image of purifed Curli. Scale bar represents 100 nm. C: Purified Curli displayed a maximum absorbance at the frequency of 1620 cm^{-1}.

The ultrastructure of purified Curli was investigated by negative-stain EM. Purified Curli displayed long thin unbranched fibrils that tend to associate laterally (**Figure 15B**). These

properities are characteristic for amyloid-like fibrils and are similar to *in vivo* Curli fibrils (**Figure 15B**). Because of sample sonication during purification, purified Curli appear to be shorter than *in vivo* Curli and CsgA fibrils (Chapman, Robinson et al. 2002).

The secondary structure of purified Curli was determined using FT-IR spectroscopy. The maximum absorbance at the frequency of 1620 cm^{-1} is characteristic for the β-sheet structure of amyloids fibrils (Hiramatsu and Kitagawa 2005). The FT-IR spectrum of purified Curli exhibited a maximum absorbance at the frequency of 1620cm^{-1} indicating well-ordered amyloid fibrils with β-sheet structure (**Figure 15C**).

4.4.Solid-state NMR analysis of native Curli

Solid-state NMR spectroscopy is the most powerful technique for acquiring structural information about amyloids fibrils. In order to illustrate native Curli structure, uniformly ^{15}N,^{13}C-labeled native Curli isolated from MC4100 bacterial cells was examined by two-dimensional ^{13}C-^{13}C solid-state NMR. The ^{13}C-^{13}C correlation spectrum of native Curli was recorded by Dr. Mumdooh Ahmed with 25 ms DARR mixing on a 21T spectrometer. The ^{13}C spectrum showed intra-residue correlations for the same spin-systems. The spectrum of native Curli showed strong cross-peaks that connect ^{13}C chemical shifts of bonded ^{13}C sites, allowing assignments to residues type based on their average chemical shifts using CcpNMR software, resulting in a total of 33 amino acid residues were identified and the number for each kind of residue was randomly marked, which included Alanine (8), Isoleucine (3), Glutamine (6), Serine (9), Threonine (6) and Valine (1) (**Figure 16**).

The ΔCα-ΔCβ value provides an insight into the relationship between chemical-shift and protein secondary structure, for example, when ΔCα-ΔCβ value is less than 0, the amino acid is involved in β-sheet seconday structure. (Fritzsching, Yang et al. 2013). Native Curli contains very less CsgB (**Figure 15A**), so the spectrum should not show the residue's cross-peak belonging to CsgB. The protein sequence of CsgA contains a total of 3 isoleucine residues - 47Ile, 69Ile and 98Ile. All these three isoleucine residues of CsgA can be found at the isolated cross peaks (**Figure 16**): 5 cross-peaks of C_{d1}-C_{g2}, C_{g2}-$C_{d1,}$ C_{g2}-C_{α}, C_{g2}-C_{β} and C_{g1}-C_{β} for the first Ile (1I), 4 cross-peaks of C_{d1}-C_{g2}, C_{g2}-$C_{d1,}$ C_{g2}-C_{α} and C_{g2}-C_{β} for the second Ile (2I) and 4 cross-peaks of C_{d1}-C_{g2}, C_{g2}-$C_{d1,}$ C_{g2}-C_{α} and C_{g2}-C_{β} for the last Ile (3I). The ΔCα-ΔCβ values of these isoleucine residues were calculated and less than 0 (**Figure 17**), the values indicated that 47Ile, 69Ile and 98Ile are involved in β-sheet segments, and it could

be postulated that the Curli adopts a sole well-defined structure. For the other 30 residues, their ΔCα-ΔCβ values were also less than 0 (**Figure 17**), suggesting that these residues should be involved in β-sheet segments as well (**Table 16; Figure 18**).

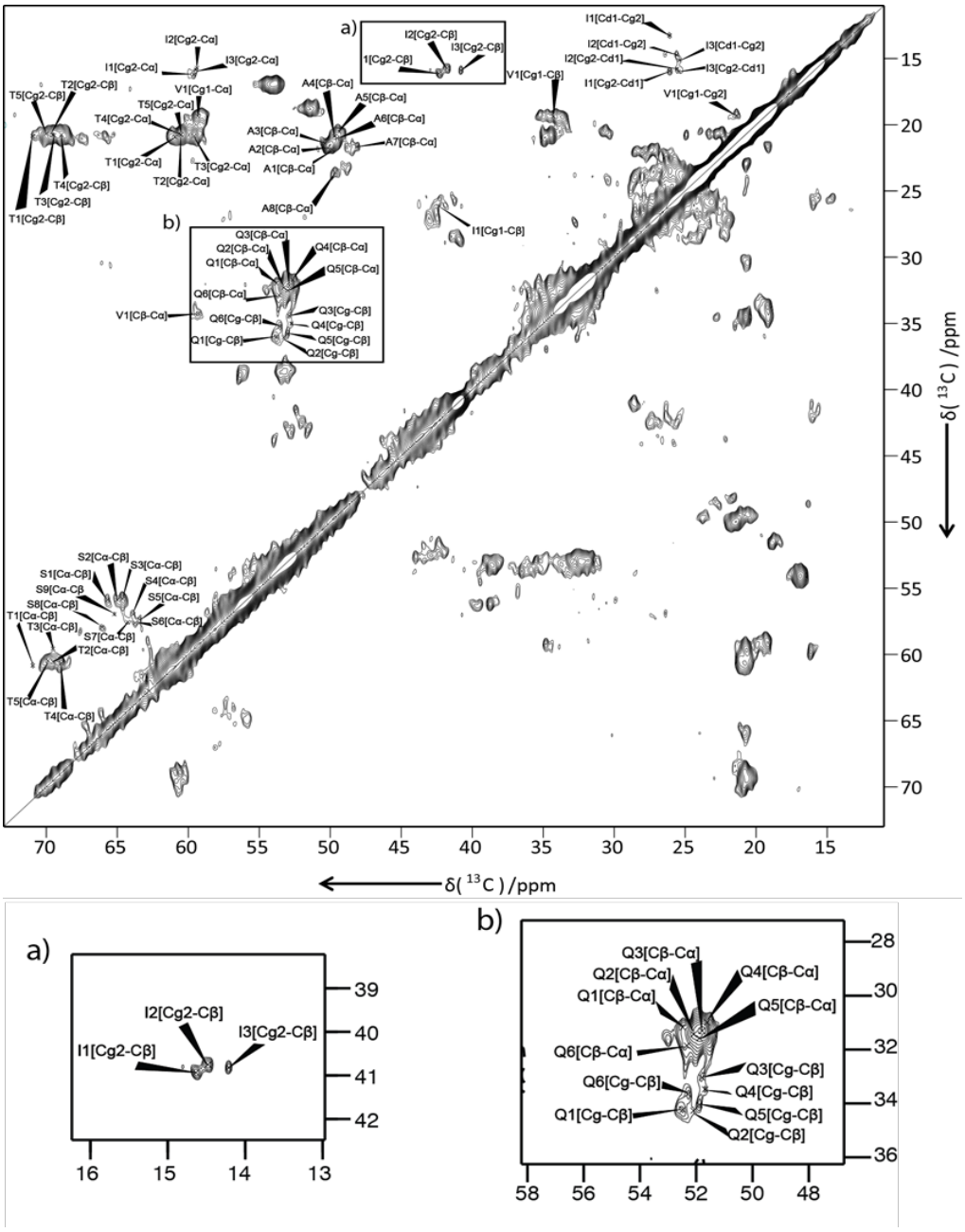


Figure 16: Two dimensional ^{13}C-^{13}C correlation spectrum of native Curli showing intraresidue correlations

Based on the average Chemical shifts of specific amino acid, a total of 33 amino acid residues were identified. The number for each kind of residue was randomly named. a) and b) are the selected cross-peaks of isoleucine (I) and glutamine(Q) respectively.

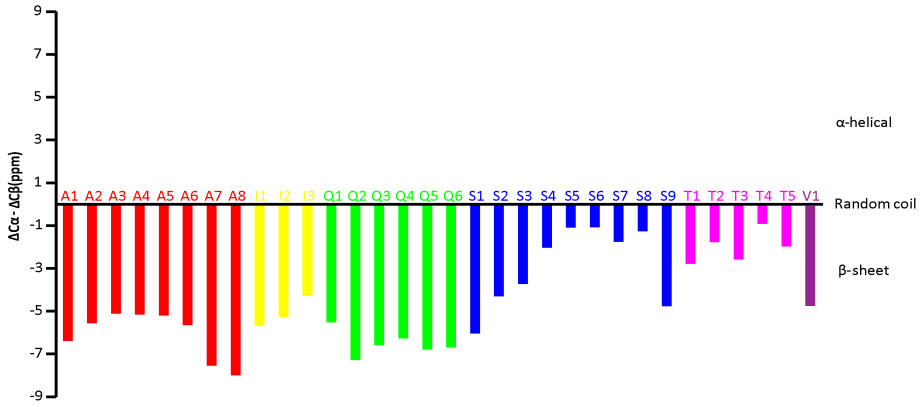


Figure 17: Secondary chemical shifts for amino acid-type assigned residues in native Curli

ΔCα-ΔCβ values (denoted secondary chemical shift) of the amino acid residues were identified. The bars represent expected values (less than 0 ppm) of ΔCα-ΔCβ for β-sheet residues, these colorful bars corresponding to the secondary chemical shift for each amino acid residue according to the classification β-sheet structures.

Table 16: Amino acid residues of native Curli were identified in the ^{13}C-^{13}C solid-state spectrum

Residue type	Number of picked peak sets (in amyloid core/total in sequence)	β-sheet secondary structure (number of picked peak sets)
A (Ala)	8 (10/10)	8
I (Ile)	3 (3/3)	3
Q (Gln)	6 (10/11)	6
S (Ser)	9 (11/12)	9
T (Thr)	6 (9/9)	6
V (Val)	1 (6/8)	1

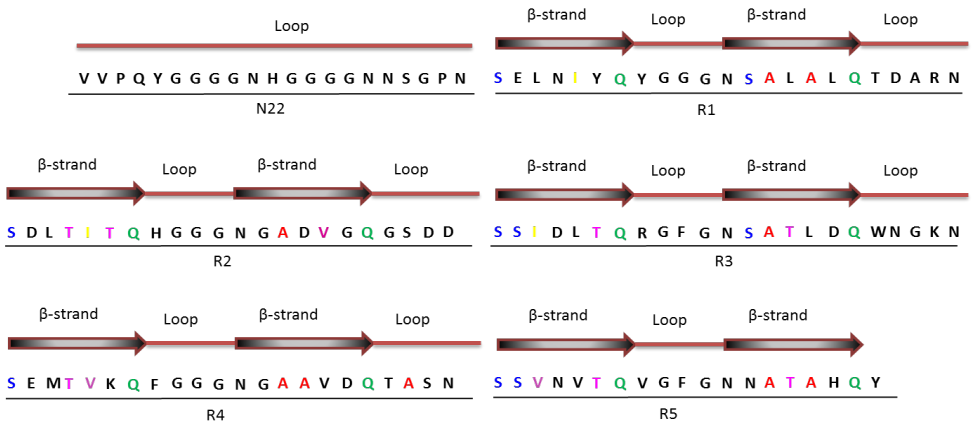


Figure 18: A putative secondary structure of native Curli (CsgA)

8 Alanine (red color), 3 Isoleucine (yellow color), 6 Glutamine (green color), 9 Serine (blue color), 6 Threonine (pink color) and 1 Valine (purple color) were identified in solid-state spectrum of native Curli, all these residues should be involved in β-sheet segments along the CsgA sequence.

4.5. Discussion

Medium ingredients strongly affect the yield of native Curli (Chapman, Robinson et al. 2002, Jubelin, Vianney et al. 2005). The results confirmed that salts and glucose are two important components which can strongly affect the yield of native Curli resulting in an optimized medium for isotope labeling that contains only 10 mM Na_2HPO_4 / KH_2PO_4 and 0.2% glucose.

The procedure of Curli purification presented here is modified from the Collinson SK's method (Collinson, Emody et al. 1991), the major difference is that Curli was shear off from bacterial surface using a sonifier water-bath, because amyloid fibrils can break up by sonication. During the processing, not only Curli fibrils but also outer membrane proteins can be sheared off (**Figure 14B**). In order to remove the contaminants and get Curli of higher purity and quality, curde Curli sample was treated with two detergents-SDS and Triton-X100, and further purified using Gel-purification as described in Collinson SK's method.

Purified native Curli displayed long thin fibrils with a β-sheet secondry structure that is similar to *in vivo* Curli fibrils (**Figure 3A & 15**). Native Curli contained approximately 90-95% CsgA and 5-10% CsgB (**Figure 15A**).

Amyloid core (residues 43-151) of CsgA is composed of five imperfect repeating units (R1-R5) and distinguished by the consensus sequence of Ser-X5-Gln-X5-Asn-X5-Gln (X represents any residue) (**Figure 2A**) (Chiti and Dobson 2006, Wang, Smith et al. 2007). A 2D ^{13}C-^{13}C solid-state NMR spectrum of native Curli was obtained on approximately 25 mg of uniformly ^{15}N,^{13}C-labeled Curli isolated from MC4100 bacterial cells. A total of 33 residues have been identified in the correlation spectrum (**Figure 16; Table 16**), and all these identified residues should be well placed in β-strand regions (**Figure 18**). For example, all three isoleucine residues of CsgA should be involved in β-sheet segments and 6 conserved glutamine residues also could be expected β-sheet residues. Combined with the FT-IR result of native Curli, these findings could explain that native Curli displays a β-sheet secondary structure.

The number of resonances observed in the spectrum of native Curli does not correlate to the number of amino acid residues of the CsgA protein. This imply that the native Curli is a great degree of resonance overlap due to sequence and structure homology adopted by different repeats (Shewmaker, McGlinchey et al. 2009).

5. CsgB interspersed between CsgA within the Curli fibrils

Curli consist of CsgA, the major subunit, and CsgB, the minor subunit. *In vitro*, freshly CsgA and CsgB exhibit intrinsically disordered properties as monomers (Wang, Smith et al. 2007). Both CsgA protein and CsgB protein individually are able to self-associate to form amyloid-like fibrils in the proper conditions (Wang, Smith et al. 2007, Dueholm, Nielsen et al. 2011). *E.coli* Curli contains both CsgA and CsgB but the contents have a 10-20 folds difference (Wang, Hammer et al. 2008). The architecture of Curli is described as the fibril that a CsgB molecule nucleates CsgA molecules for elongating perpendicular to the elongating axis. There are two proposed model of Curli: the model A shows that the CsgB is located in the ends of Curli fibrils, the model B shows that CsgB is not only in the ends but also in between CsgA molecules interacting with neighbouring CsgA molecules in Curli (**Figure 20**).

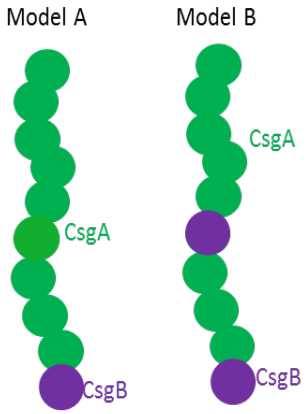


Figure 20: Two proposed models of Curli

Model A: CsgB is located in the ends of Curli; Model B: CsgB is not only in the ends but also in between CsgA molecules in Curli.

In the following studies, mixed fibrils resulting from the mixtures of CsgA protein and CsgB protein will be studied using controlled shear-forces combined with seeding assays (**5.1**), and Immunogold labelling TEM (**5.2**) as well as H/D exchange NMR technologies (**5.3**) to reveal how CsgA and CsgB are assembled into mixed fibrils *in vitro*.

5.1. Analysis of mixed CsgA-CsgB fibrils by controlled Shear-force combined with seeding assays

Mixed CsgA-CsgB fibrils can form from the mixtures of CsgA protein and CsgB protein. CsgA fibrils and mixed fibrils are processed by controlled shear-force at various conditions resulting in smaller fibril fragments. Freshly buffer-exchanged CsgA is seeded by these processed fibrils and their seeding capacities can be evaluated by ThT fluorescence.

The main aim of this part is how different strengths and cycles affect the sending capacity of processed fibrils.

5.1.1. Purification of recombinant CsgA and CsgB proteins

The recombinant CsgA and CsgB proteins were expressed in *E.coli* T7 cells (**3.2.2**). Both recombinant CsgA and CsgB proteins were affinity purified under denaturing condition because they were overexpressed in inclusion bodies as illustrated in **3.2.2**. Approximately 2-5 mg of pure CsgA and CsgB could be purified from one liter of LB medium.

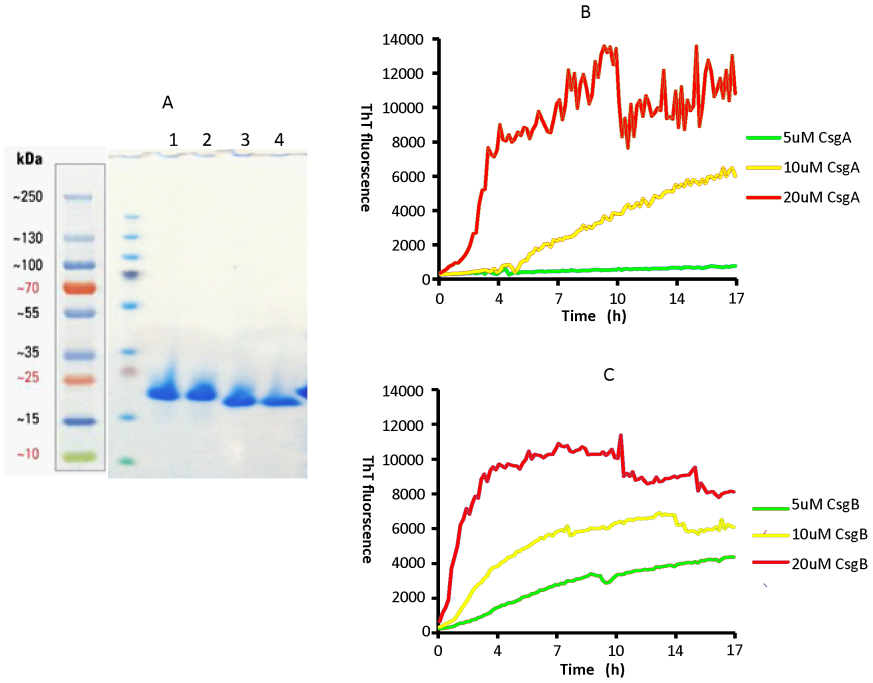


Figure 21: Biochemical properties of CsgA and CsgB protein *in vitro*

A: Coomassie stained SDS-PAGE gel of recombinant CsgA (lane 1 & 2) and CsgB (lane 3 & 4). Pre-stained protein ladder is indicated on the left.
B & C: Self-aggregation of CsgA (B) and CsgB (D) in pH 7.4 buffer were followed by ThT fluorescence. Increase of ThT fluorescence is depended on incubation time (0-17 h) and protein concentrations (5 µM, 10 µM and 20 µM).

Both CsgA and CsgB proteins were stored in denaturing buffers, because they are extremely unstable as well as aggregating rapidly (f.g. CsgA aggregation, approximately 2 h in pH 7 buffer) in the conditions without denaturants such as GuHCl (Chapman, Robinson et al. 2002). Purified CsgA (lane 1 & 2) and CsgB (lane 3 & 4) were analyzed by SDS-PAGE. The molecular weight of CsgA protein is 14 kDa, however, it runs at around 18 kDa; the molecular weight of CsgB protein is 14.5 kDa, it also shows around 18 kDa (**Figure 21A**).

5.1.2. CsgA fibrils and CsgB fibrils have similar biophyscical properties

The amyloidogenic properties of CsgA and CsgB proteins were analyzed using the amyloid specific dye-thioflavin T (ThT). The aggregation kinetics was followed by the ThT fluorescence using a Plate-reader (**3.2.3**).

Denatured CsgA protein was buffer-exchanged into pH 7.4 buffer removing the denaturants, 5 μM, 10 μM and 20 μM of buffer-exchanged CsgA protein were mixed with ThT dye respectively and measured immediately. The ThT binding showed increasing fluorescence over time in a CsgA concentration manner. Compared to the self-aggregation kinetics of 5 μM CsgA sample, the lag time of 10 μM CsgA was 60 mins and the lag time of 20 μM CsgA was 280 mins (**Figure 21B**). The lag time of CsgA protein is not only depended on the protein concentration, but also strongly depended on buffers (Chapman, Robinson et al. 2002, Dueholm, Nielsen et al. 2011).

Unlike CsgA protein, CsgB protein has to be buffer-exchanged into pH 3 buffer, because CsgB protein aggregates fastly pH 7.4 buffer (Madhu Nagaraj, PhD thesis). Buffer pH was switched from 3 to 7.4 when CsgB protein was mixed with ThT dye. Increasing of fluorescence over time is also depended on CsgB protein concentration (**Figure 21C**). CsgB protein showed faster aggregation kinetics (5 μM, 10 μM and 20 μM) compared to CsgA protein.

The ultrastructure of CsgA fibrils and CsgB fibrils was investigated by TEM. 20 μM of freshly buffer-exchanged CsgA protein and CsgB protein were immediately used to form fibrils in pH 7.4 buffer. Negative-stained TEM images of CsgA fibrils and CsgB fibrils revealed amyloid-like fibrils similar to native Curli (**Figure 22**). Both CsgA fibrils and CsgB fibrils exhibited large clumps.

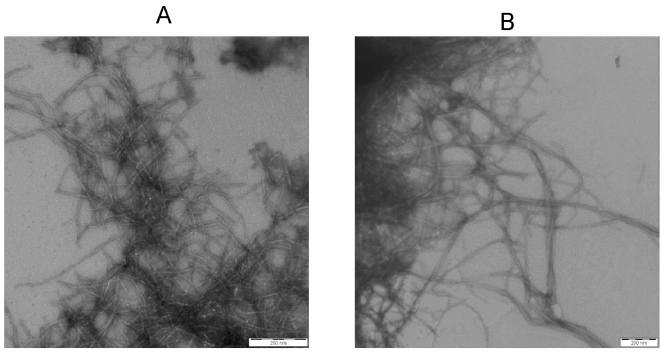


Figure 22: Negative-stained TEM images of CsgA fibrils and CsgB fibrils

Negative-stained TEM images of CsgA fibrils (A) and CsgB fibrils (B) formed in pH 7.4 buffer, scale bar represents 200 nm.

5.1.3. CsgA aggregation kinetics in pH 7.4 buffer and pH 9 buffer

Buffer pH is a crucial factor to affect CsgA aggregation kinetics (Dueholm, Nielsen et al. 2011). CsgA aggregation kinetics in pH 7.4 buffer and pH 9 buffer respectively were followed by the ThT fluorescence. Compared with CsgA aggregation in pH 7.4 buffer, CsgA aggregation kinetics in pH 9 buffer were significantly slowed down. The lag time of 20 μM CsgA in pH 7.4 buffer was approximately 1.5 h (**Figure 23A**), but the lag time in pH 9 buffer increased markedly greater than 8-fold up to about 15 h (**Figure 23B**). The lag phase of 10 μM CsgA also showed a similar increase. Interestingly, except the lag phase increased along with buffer pH, the ThT fluorescence value in stationary phase also increased significantly. This could be due to either genuine modification in CsgA fibril structure or pH dependence of ThT dye binding.

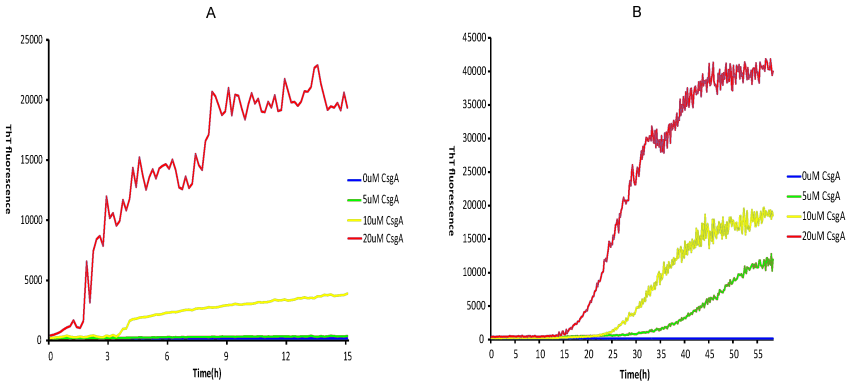


Figure 23: Self-aggregation of CsgA protein was in pH 7.4 buffer and pH 9 buffer respectivey

CsgA was self-aggregated (0, 5, 10 and 20 μM) in pH 7.4 buffer (A) and pH 9 buffer (B) respectively followed by the ThT fluorescence.

5.1.4. CsgA fibrils have the similar secondary structure formed in pH 7.4 buffer and pH 9 buffer

20 µM of freshly buffer-exchanged CsgA protein was used to form amyloid fibrils in pH 7.4 buffer and pH 9 buffer respectively without shaking at RT for 24 hours (**3.2.2**). The morphology of CsgA fibrils was examined by negative-stained TEM. These CsgA fibrils formed in different buffer exhibited similar amyloid-like fibrils. CsgA fibrils formed in pH 7.4 buffer showed a lot of large clumps (**Figure 24A**); however, CsgA fibrils formed in pH 9 buffer displayed much less clumps (**Figure 24B**).

The secondary structure of the CsgA fibrils was analyzed by FT-IR spectroscopy. FI-TR exhibited very similar spectra for the CsgA fibrils from both conditions with a well-defined peak at 1624 cm^{-1} (**Figure 24C**), indicating β-sheet structure of amyloid-like conformation (Kong and Yu 2007). CsgA fibrils formed in both pH buffers have quite similar β-sheet secondary structure.

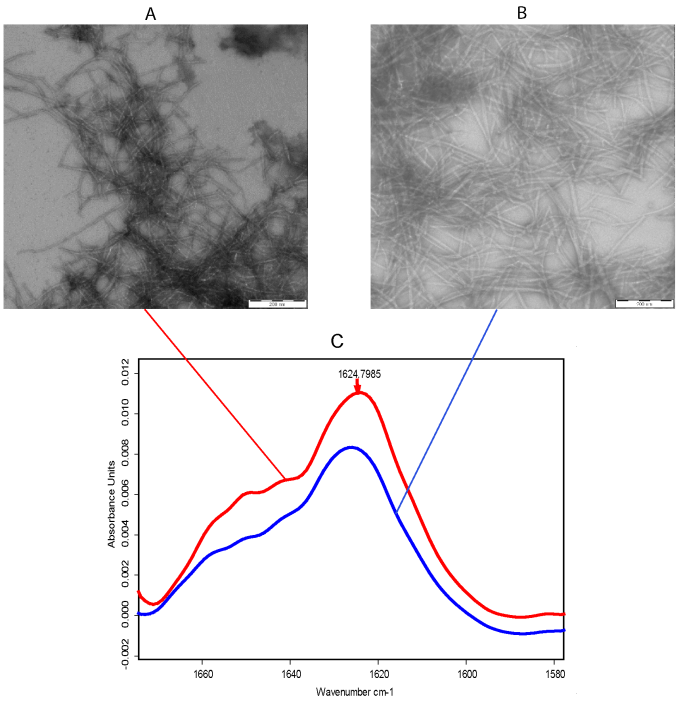


Figure 24: CsgA fibrils formed in different buffers indicated the similar secondary structure

Negative-stained TEM images of CsgA fibrils formed in pH 7.4 buffer (A) and pH 9 buffer (B), scale bar represents 200 nm; FTIR spectra of CsgA fibrils formed in pH 7.4 buffer (Red) and pH 9 buffer (Blue) exhibited similar secondary structure.

5.1.5. Mixed CsgA-CsgB fibrils formed from co-aggregation CsgA protein and CsgB protein

In Curli assembly, the minor CsgB subunit is a potent nucleator, on the bacterial surface. It is believed that unfolded CsgA monomers fold into proper conformation upon association with folded CsgB monomer, forming nucleation centers (Chapman, Robinson et al. 2002). CsgB protein can also nucleate CsgA protein *in vitro* (Madhu Nagaraj, PhD thesis).

CsgA protein was buffer-exchanged into pH 7.4 buffer and CsgB protein was buffer-exchanged into pH 3 buffer. All six protein samples of 10 µM CsgA alone, 10 µM CsgB alone and four mixtures of 9.5 µM CsgA : 0.5 µM CsgB, 9.0 µM CsgA : 1.0 µM CsgB, 8.0 µM CsgA : 2.0 µM CsgB, 5.0 µM CsgA : 5.0 µM CsgB were immediately prepared in the presence of ThT dye and their aggregations kinetics were measured by ThT fluorescence. The lag phase of aggregation was dramatically shortened in a concentration dependent manner of the CsgB (**Figure 25**), even the concentration of CsgB is 5%. The lag phase of CsgB alone was similar to the lag phases of 8.0 µM CsgA : 2.0 µM CsgB, 5.0 µM CsgA : 5.0 µM CsgB.

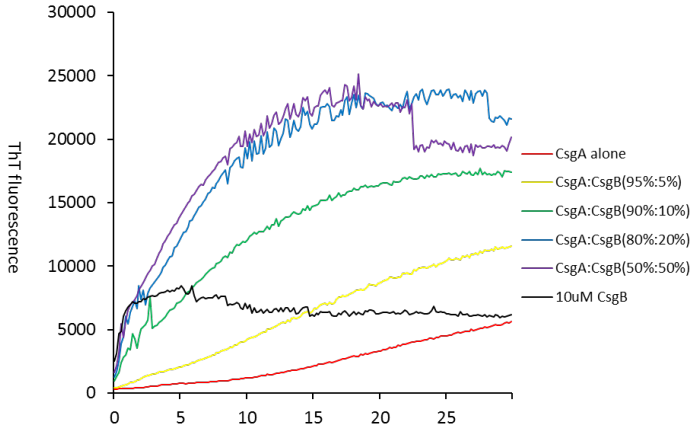


Figure 25: Co-aggregation of CsgA protein and various amount of CsgB protein followed by ThT fluorescence

CsgA protein and various amount of CsgB protein (0, 5, 10, 20 and 50%) were co-aggregated and followed by ThT fluorescence.

Fibrils were formed from 10 µM CsgA alone, 10 µM CsgB alone and four mixtures of 9.5 µM CsgA : 0.5 µM CsgB, 9.0 µM CsgA : 1.0 µM CsgB, 8.0 µM CsgA : 2.0 µM CsgB, 5.0 µM CsgA : 5.0 µM CsgB in pH 7.4 without shaking at RT for 2 days, and analyzed by negative-stained TEM. All samples exhibited similar amyloid-like fibrils with large clumps or dense amorphous aggregates (**Figure 26**).The results revealed that the mixed fibrils even in presence

of 50% CsgB showed similar fibrils as CsgA fibrils or CsgB fibrils. These kinds of mixed fibrils containing various amount of CsgB protein can be used the following experiments.

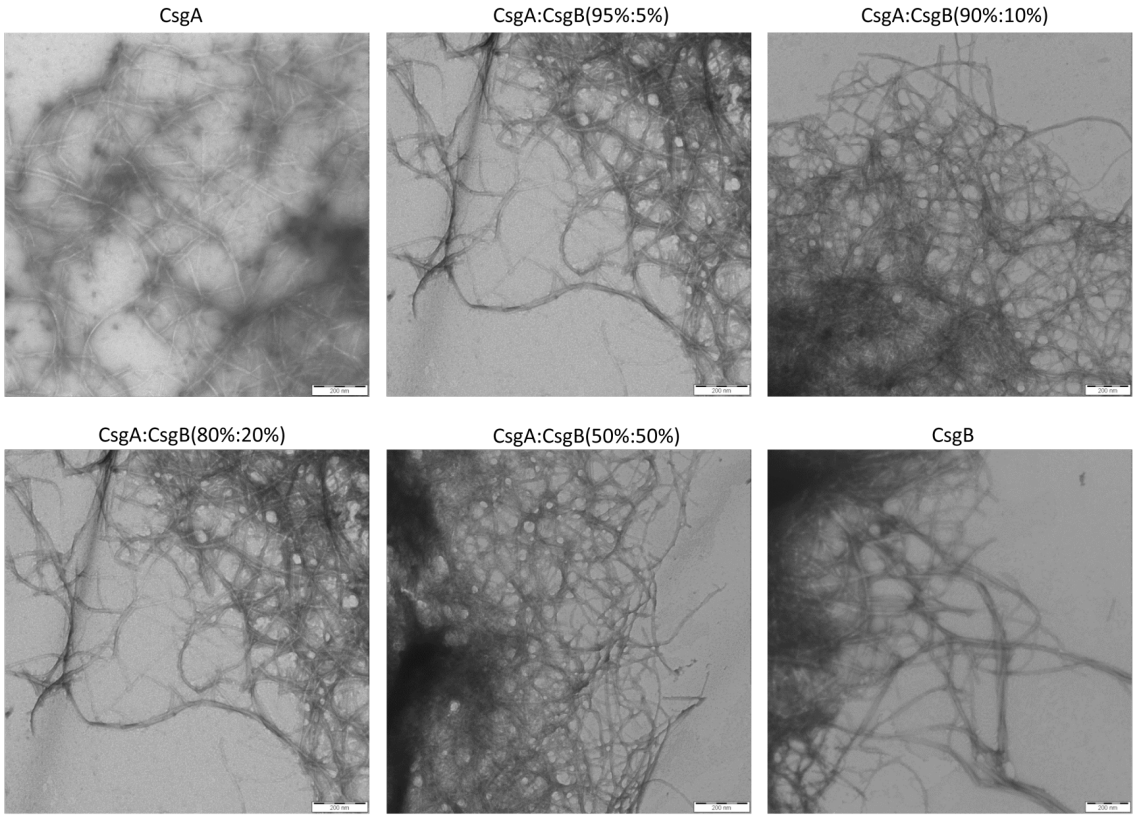


Figure 26: Co-aggregation of CsgA protein by CsgB protein monitored by negative-stained TEM

Aggregation of CsgA protein in presence of various amount of CsgB protein (0, 5, 10, 20 and 50%) monitored by negative-stain TEM.

5.1.6. Effect of CsgA fibrils concentration on the seeding of CsgA aggregation in two buffers

CsgA protein was freshly buffer-exchanged into pH 7.4 buffer and pH 9 buffer respectively. 10 μM of monomeric CsgA protein was mixed with 1%, 2%, 4% and 8% of 2-day old CsgA fibrils respectively in presence of ThT dye and further immediately measure the ThT fluorescence.

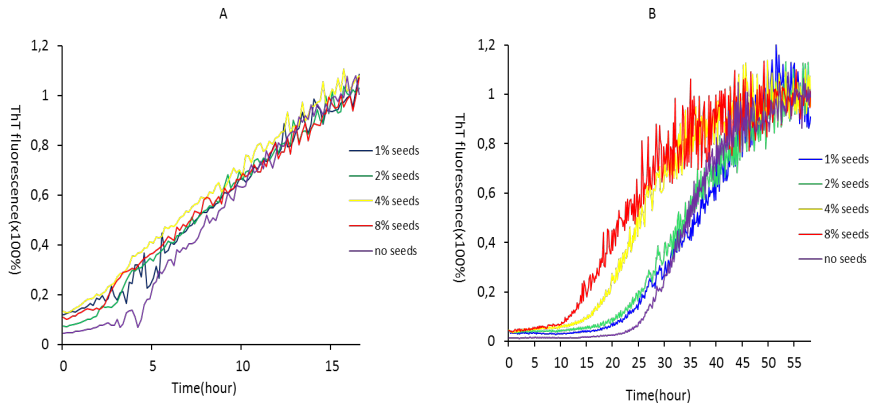


Figure 27: CsgA protein was seeded by the seed of CsgA fibrils in pH 7.4 and pH 9 respectively

10 μM of CsgA protein was seeded by 1%, 2%, 4% and 8% of 2-day old CsgA fibrils respectively followed by ThT fluorescence. A: the reaction in pH 7.4 buffer; B: the reaction in pH 9.0 buffer

In pH 7.4 buffer, the lag phases decreased slightly when CsgA were incubated with various amount of CsgA fibrils. And it was difficult to compare with the lag times. However, In pH 9 buffer, the lag phases decreased markedly with an increasing CsgA fibrils seed concentration as well as the lag phases were highly depended on the concentration (0-8%), even the lowest seed concentration (1%) had a clear effect on CsgA aggregation (**Figure 27B**).

The reason could be that the process of CsgA aggregation is faster than the process of CsgA seeded by CsgA fibrils in pH 7.4 buffer, so the CsgA fibrils seed has less effect on the reaction. In the following experiments, CsgA aggregation was in pH 9 buffer.

5.1.7. Effect of processed CsgA fibrils on the aggregation of CsgA protein

As mentioned above, the lag phase significantly decreased with increasing CsgA fibrils seed concentration in pH 9 buffer. Each fibril used as a single end, monomer CsgA adds to this end while start a new CsgA fibril formation. The more of the ends, the more of new CsgA fibrils can be formed.

The Selective Shearing Amplification array is an apparatus that can break amyloid fibrils with defined shear forces generated via a specific rotational speed of the rotor. In this experiment, 2-day old CsgA fibrils was formed in pH 7.4 buffer and further treated with 1% SDS to remove any non-specific aggregates. The fibrils were subjected to various cycles (3, 6 and 12 cycles) of 5 s rotation and 30 s resting phase. The rotation frequencies were set to 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively. These processed fibrils were immediately used at a concentration of 3% to seed freshly buffer-exchanged CsgA protein in pH 9 buffer.

The effect of differently processed fibrils on the kinetics of CsgA aggregation was investigated (**Figure 28A**). Two of lag phases were significantly decreased (CsgA were seeded by 760Hz and 960Hz processed fibrils). The other 6 of lag phases (CsgA were seeded by 240Hz, 300Hz, 380Hz, 480Hz, 600Hz and unprocessed fibrils) were at the similar levels. This result indicated that 760Hz could be the critical frequency to break CsgA fibrils, and below this frequency such as 600Hz, CsgA fibrils could not be broken up.

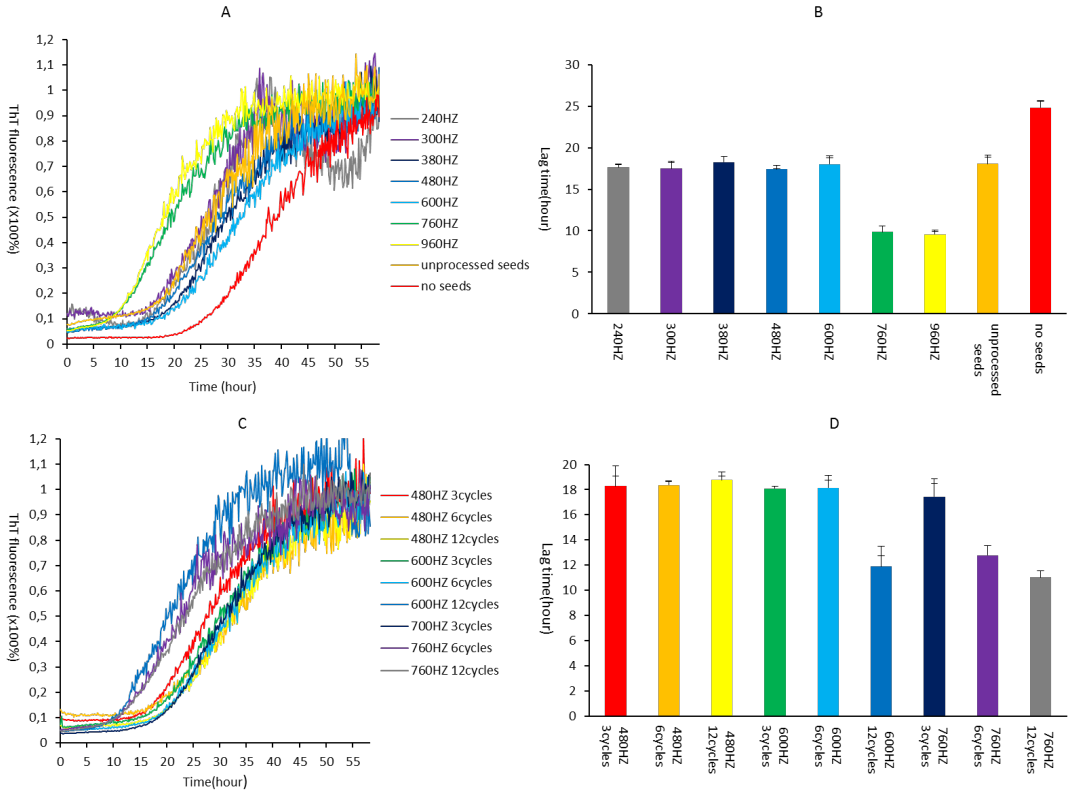


Figure 28: Effect of processed CsgA fibrils on the aggregation of CsgA protein

CsgA fibrils (2-day old) were processed by Selective Shearing Amplification array with 3, 6 and 12 cycles of 5 s rotation at 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively and 30s resting phase without rotation. 3% of the processed fibrils was used to seed freshly buffer-exchanged monomer CsgA in pH 9 buffer in presence of ThT dye. Each experiment was replicated three times. The lag phase was calculated.
The fibrils were processed by 6 cycles (A) and their lag phases (B).
The fibrils were processed by for 3, 6, 12 cycles respectively at 480Hz, 600Hz and 760Hz (C) and their lag phases (D).

In order to test what the cycle numbers affect CsgA aggregation, CsgA fibrils were processed by 3, 6, 12 cycles respectively at 480Hz, 600Hz or 760Hz. Obviously, CsgA fibrils weren't broke up at the frequencies of 480Hz and 600Hz for 6 cycles or 3 cycles as well as 760Hz for 3 cycles. However, CsgA fibrils were broken up at the frequency of 600Hz for 12 cycles (**Figure 28B**). Therefore it proves that both the frequency and the cycle number are two important factors for the break-up of CsgA fibrils processed by Selective Shearing

Amplification array.

5.1.8. Effect of processed mixed CsgA-CsgB fibrils on the aggregation of CsgA protein

Three mixed CsgA-CsgB fibrils (2-day old), which contained 5%, 10% and 20% of CsgB protein respectively, were produced in pH 7.4 buffer. Mixed fibrils were subjected to various cycles (3, 6 and 12 cycles) of 5 s rotation and 30 s resting phase. The rotation frequencies were controlled by 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively. These processed mixed fibrils were immediately used at a concentration of 3% to seed freshly buffer-exchanged CsgA protein in pH 9 buffer.

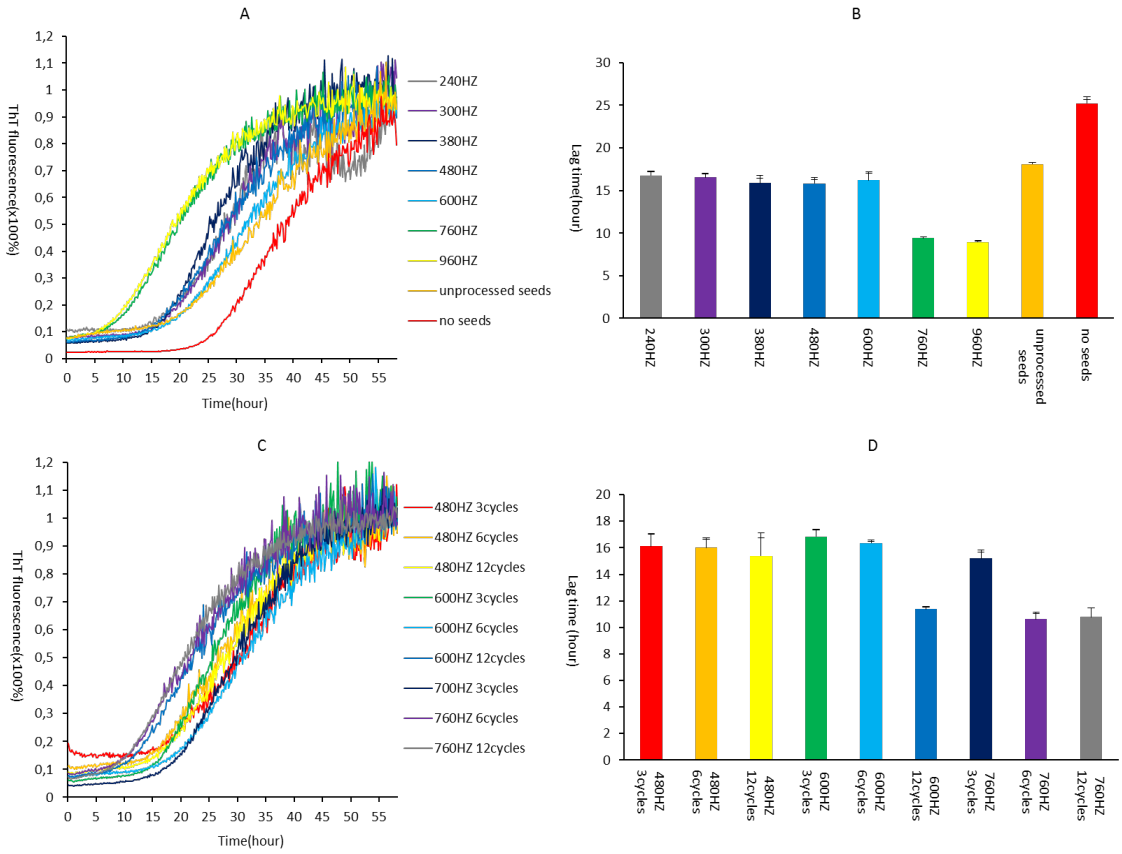


Figure 29: Effect of processed mixed CsgA-CsgB (95%:5%) fibrils on the aggregation of CsgA protein

Mixed CsgA-CsgB (95%:5%) fibrils (2-day old) were processed by Selective Shearing Amplification array with 3, 6 and 12 cycles of 5 s rotation at 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively and 30 s resting phase without rotation.3% of the processed fibrils was used to seed freshly buffer-exchanged monomer CsgA in pH 9 buffer in presence of ThT dye. Each experiment was replicated three times.The lag phase was calculated.
The fibrils were processed by 6 cycles (A) and their lag phases (B).
The fibrils were processed by for 3, 6, 12 cycles respectively at 480Hz, 600Hz and 760Hz (C) and their lag phases (D).

The results of mixed CsgA-CsgB (95%:5%) fibrils were similar to the previous results of

CsgA fibrils. In 6 cycles, mixed CsgA-CsgB (95%:5%) fibrils were only broken up at the critical frequency of 760Hz and the higher 960Hz (**Figure 29A**). For the further test, CsgA were seeded by mixed CsgA-CsgB (95%:5%) processed by different cycles (**Figure 29B**), the lag phase linked to that CsgA-CsgB (95%:5%) wasn't broken up at the frequency of 600Hz for 6 cycles.

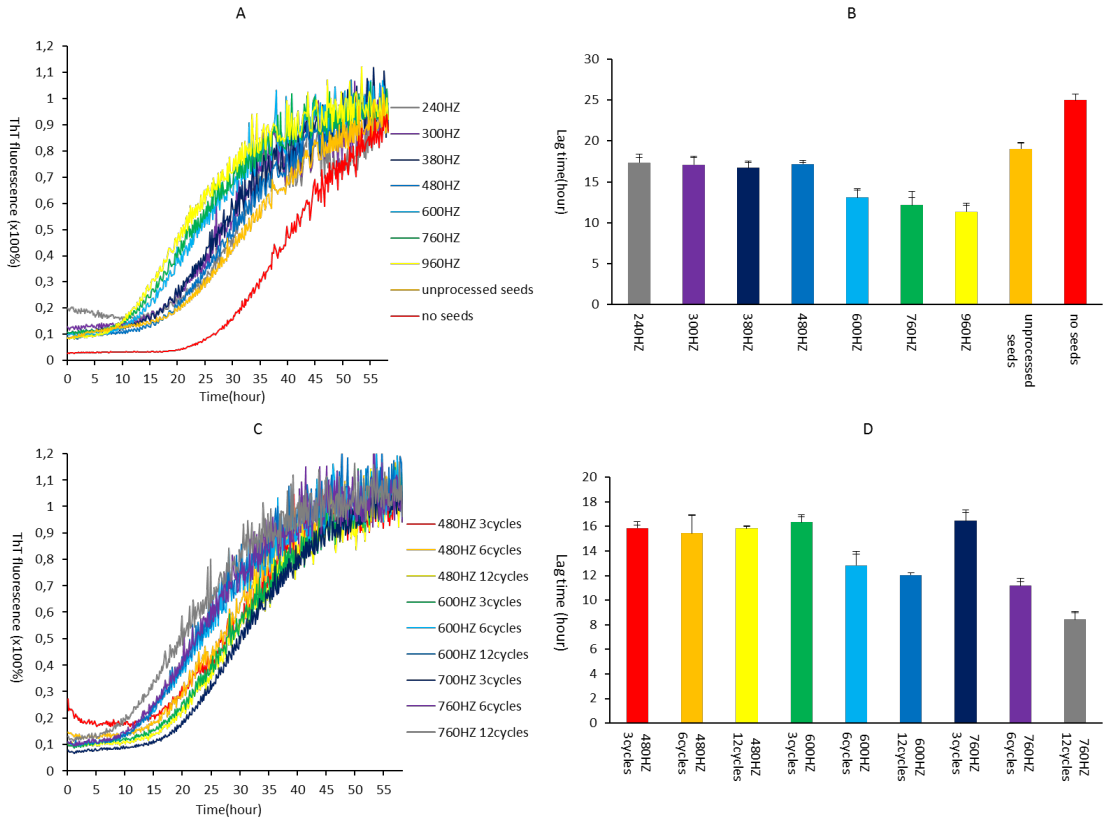


Figure 30: Effect of processed mixed CsgA-CsgB (90%: 10%) fibrils on the aggregation of CsgA protein

Mixed CsgA-CsgB (90%:10%) fibrils (2-day old) were processed by Selective Shearing Amplification array with 3, 6 and 12 cycles of 5 s rotation at 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively and 30 s resting phase without rotation. 3% of the processed fibrils was used to seed freshly buffer-exchanged monomer CsgA in pH 9 buffer in presence of ThT dye. Each experiment was replicated three times. The lag phase was calculated.
The fibrils were processed by 6 cycles (A) and their lag phases (B).
The fibrils were processed by for 3, 6, 12 cycles respectively at 480Hz, 600Hz and 760Hz (C) and their lag phases (D).

Next, mixed fibrils containing a larger proportion of CsgB were analyzed the same way. As the previous results of CsgA fibrils and mixed CsgA-CsgB (95%:5%) fibrils in 6 cycles, the lag phases linked to that mixed CsgA-CsgB (90%:10%) fibrils were broken up at the frequencies of 760Hz and 960Hz as well as weren't at the frequencies from 240-480Hz (**Figure 30A**). The lag phase was significant decreased because the mixed fibrils can be

broken up at the frequency of 600Hz, and it was more importantly that the additional ends were distinguished. The further results strengthen this view, at the frequency of 600Hz; mixed CsgA-CsgB (90%:10%) fibrils were broken up for both 6 cycles and 12 cycles (**Figure 30A**).

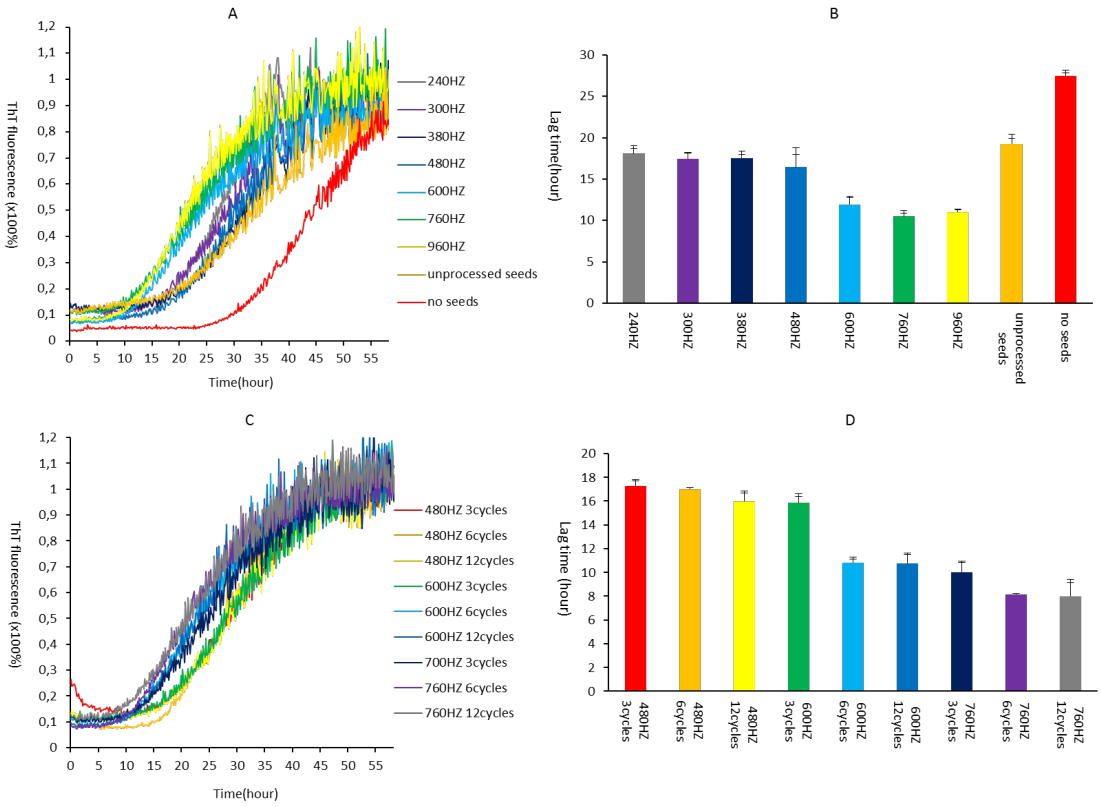


Figure 31: Effect of processed mixed CsgA-CsgB (80%:20%) fibrils on the aggregation of CsgA protein

Mixed CsgA-CsgB (80%:20%) fibrils (2-day old) were processed by Selective Shearing Amplification array with 3, 6 and 12 cycles of 5 s rotation at 240Hz, 300Hz, 380Hz, 480Hz, 600Hz, 760Hz and 960Hz respectively and 30 s resting phase without rotation. 3% of the processed fibrils was used to seed freshly buffer-exchanged monomer CsgA in pH 9 buffer in presence of ThT dye. Each experiment was replicated three times. The lag phase was calculated.
The fibrils were processed by 6 cycles (A) and their lag phases (B).
The fibrils were processed by for 3, 6, 12 cycles respectively of at 480Hz, 600Hz and 760Hz (C) and their lag phases (D).

The last experiment was related to the mixed CsgA-CsgB (80%:20%) fibrils containing 20% of CsgB protein. For 6 cycles, the lag phase only significantly reduced at frequency of 600Hz, 760Hz and 900Hz (**Figure 31A**), and these results were duplicated in the further experiment. Interestingly, owing to 10% more in mixed CsgA-CsgB (80%:20%) fibrils, the lag phase (760Hz, 3 cylces) significantly decreased for the first time (**Figure 31B**).

5.2. Analysis of mixed CsgA-CsgB fibrils by Immunogold labelling TEM

In this part, the cysteine-CsgB protein was produced, and then mixd with CsgA protein to form the mixed CsgA-CsgB fibrils. Cysteine-CsgB can be specific biotinylated by a chemical reagent. Gold-labelled biotin antibody is used to detect biotinylated CsgB and the mixed fibrils are further investigated by TEM.

5.2.1. Expression and purification of cysteine-CsgB protein

The cysteine-CsgB construct was expressed in *E.coli* T7 cells and was further affinity purified in denaturing condition (**3.2.2**). Approximate 2 mg of yield could be purified from one liter of LB medium. The molecular weight of cysteine-CsgB is 15 kDa, but it runs at approximately 18 kDa (**Figure 32A**).

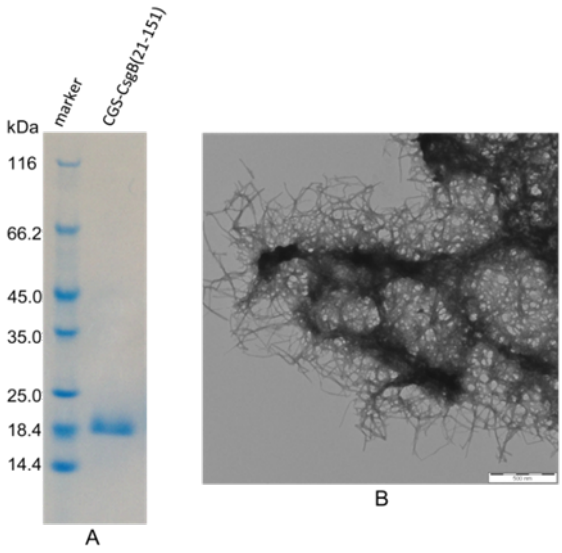


Figure 32: Characterization of cysteine-CsgB protein

A: Coomassie stain SDS-PAGE of cysteine-CsgB protein. Molecular weight of cysteine-CsgB is 15 kDa, however, it runs at around 18.4 kDa. Protein ladder is indicated on the left.
B: Negative-stained TEM image of cysteine- CsgB fibrils formed in pH 7.4 buffer. Scale bar represents 200 nm.

5.2.2. Cysteine-CsgB protein forms amyloid-like fibrils *in vitro*

The ultrastructure of cysteine-CsgB fibrils was investigated by negative-stained TEM (**3.2.3**). 20 μM of freshly buffer-exchanged cysteine-CsgB was immediately used to form amyloid-like fibrils in pH 7.4 buffer. Negative-stained TEM image of cysteine-CsgB fibrils revealed

amyloid-like fibrils similar to CsgA fibrils and CsgB fibrils, and the fibrils formed the large clumps as well (**Figure 32**).

5.2.3. Biotinylation of cysteine-modified CsgB protein

Cysteine residues in cysteine-CsgB protein were reduced by the disulfide-reducing agent TCEP (**3.2.4**). And further reduced cysteine-CsgB protein was buffer-exchanged into denaturing buffer (8 M Guandinine-HCl, 100 mM Kpi, pH 7.4 buffer) to remove TCEP. Reduced cysteine-CsgB protein reacted with the EZ-Link Maleimide PEG2-Biotin (Thermo) reagent. After biotinylation, the biotinylated protein was buffer-exchanged into pH 3 buffer and the molecular weight was measured immediately using MALDI-MS. MALDI-MS showed a 14.763 kDa of Cysteine-CsgB protein (**Figure 33A**) and a 15.291 kDa of biotinylated CsgB protein (**Figure 33B**), The results indicated that almost 100% of cysteine-CsgB protein was biotinylated by the biotin reagent.

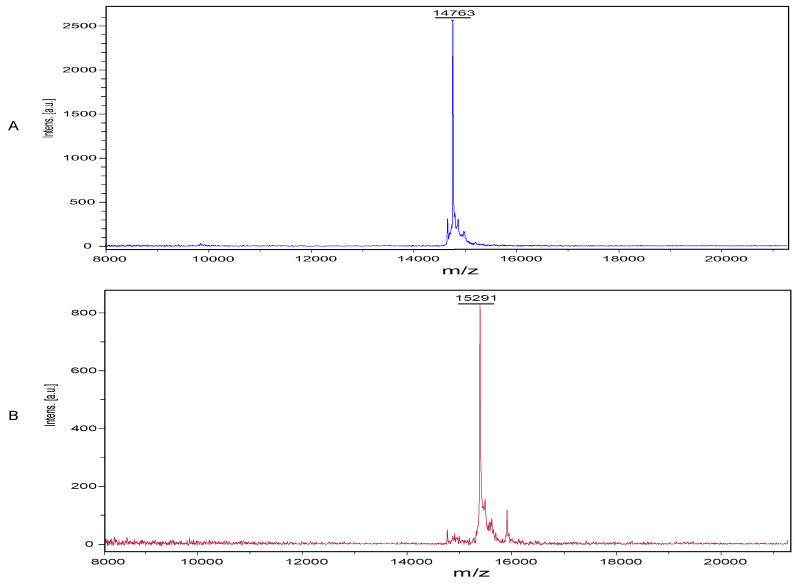


Figure 33: MALDI-MS results of cysteine-CsgB protein and biotinylated CsgB protein

MALDI-MS showed a 14.763 kDa of cysteine-CsgB protein (A) and a 15.291 kDa of biotinylated CsgB protein (B).

5.2.4. Immunogold labelling TEM analysis of biotinylated mixed CsgA-CsgB fibrils

Mixed fibrils were produced from the mixtures of CsgA protein (80%) and biotinylated CsgB protein (20%) in pH 7.4 buffer as described in **3.2.2**.

For the binding of biotinylated mixed fibrils to gold-labelled anti-biotin antibody, 100 µl of mixed fibrils (2-day old; 20 µM) was incubated with 2-4 µl of gold-labelled anti-biotin antibody (size of gold particle: 15 nm**)** for an hour at RT and unbound anti-biotin antibody was washed away using pH 7.4 buffer. CsgA fibrils and biotinylated CsgB fibrils were processed the same way and used as the controls. The detailed procedure is described in **3.2.3**.

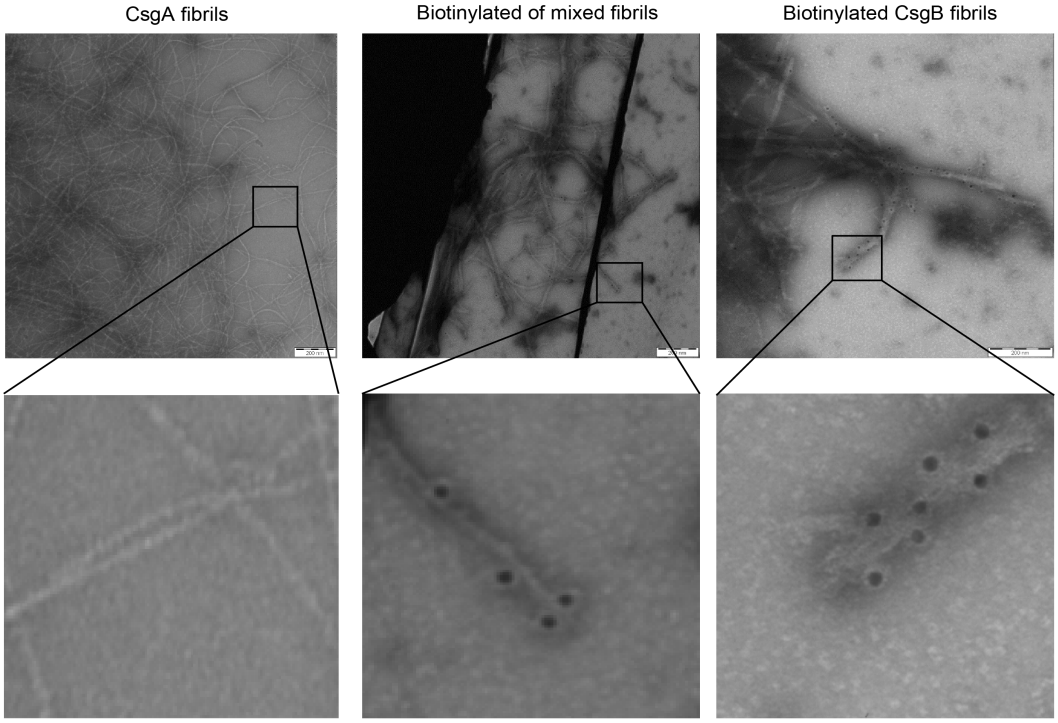


Figure 34: TEM images of gold-labeled fibrils

TEM images of the gold-labeled biotinylated mixed fibrils, the gold-labeled of biotinylated CsgB fibrils (positive control) and the gold-labeled CsgA fibrils (negative control), scale bar presents 200 nm.

TEM images of the gold-labelled of biotinylated of mixed fibrils presented that the nanogold particles is not only placed the ends but also in between CsgA molecules in the fibrils with a certain intervals, and the magnified image displayed the intervals were significantly greater than the size of more than two CsgA or CsgB molecules; for the TEM images of positive control, nanogold particles were bound to all over the CsgB fibrils with a small interval, and in such a small interval there could not accommodate a CsgB molecule if all biotinylated-CsgB molecules might be recognized by gold-labelled antibodies ; no nanogold particle was investigated in the negative control-CsgA fibrils because CsgA fibrils didn't contain biotinylated CsgB protein (**Figure 34**).

5.3. Analysis of mixed CsgA-^{15N}CsgB fibrils by H/D exchange NMR

H/D exchange NMR spectroscopy is a unique technique to study stability and thermodynamic behaviors of amyloid fibrils. Recently, four CsgB truncation mutants have been studied by H/D exchange NMR, N-terminus (21-44) of CsgB shows a single β-strand, and the C-terminus (133-151) also shows a β-strand with high-level of heterogeneity, two of these regions could provide two individual interfaces during Curli assembly such as the nucleation of CsgA (Madhu Nagaraj, PhD thesis).

In this part, mixed CsgA-^{15N}CsgB fibrils were detected by H/D exchange NMR, and the aim of this part focuses on the conformational changes of CsgB in between CsgB fibrils and mixed CsgA-CsgB fibrils.

5.3.1. Formation of mixed CsgA-^{15N}CsgB fibrils

^{15}N-labelled CsgB protein was expressed in minimal medium supplemented with ^{15}N-NH_4Cl, and then purified as described in 3.2.2.. For formation of mixed CsgA-^{15N}CsgB fibrils, denaturing CsgA protein and denaturing ^{15}N-labelled CsgB protein were buffer-exchanged into pH 7.4 buffer and pH 9 buffer respectively to remove the denaturants, and then fibrillised by increasing the buffer pH to 7.4. The mixed fibrils contained 80% of CsgA and 20% of ^{15}N-labelled CsgB. The EM images of this mixed fibrils showed good fibrils (**Figure 26 &34**) and it contained a suitable amount of CsgB for H/D exchange NMR analysis.

Most importantly, these fibrils have to be washed with 1% SDS buffer remove non-specific oligomers or aggregates (Madhu Nagaraj, PhD thesis). This step can enhance the population of ordered fibrils.

5.3.2. Backbone NH-assignments of CsgB

Dr. Madhu Nagaraj made the assigment of CsgB. Briefly, ^{15}N,^{13}C-labeled CsgB fibrils were solubilized in a mixture of DMSO and 0.1% TFA enabling complete monomerisation of the fibrils, further recorded 2D-HSQC spectra and 3D spectra of HNCA, HN (CO) CA, HNCACB and NOESY. These spectra were processed and analyzed using the programs PROSA and CARA. 92 % of CsgB resonances were assigned to the CsgB sequence (Madhu Nagaraj, PhD thesis) (**Supplementary figure 1**).

5.3.3. H/D exchange NMR analysis of CsgB in CsgB fibrils

H/D exchange allows the identification of solvent-protected backbone amide-protons indicative of H-bond formation. The HSQC spectrum (**Figure 6**) shows the reference ^{1}H-^{15}N correlation of CsgB in DMSO/0.1% TFA solvent corresponding to fully protonated CsgB. The spectrum contains one ^{1}H-^{15}N cross-peak for each backbone amide proton, allowing a residue-specific determination of the amide-proton exchange rate.

^{15}N-labelled CsgB fibrils were incubated in D_2O buffer for up to 6 w (0 h, 0.5 h, 1.5 h, 8 h, 26 h, 1 w, 2 w and 6 w) and measured by H/D exchange NMR (these works were made by Dr. Madhu Nagaraj) (Madhu Nagaraj, PhD thesis). The spectra were reprocessed by TopSpin. The H/D data were reanalyzed by CcpNMR and Metlab (with the help of Dr. Gang Zhao). After 1.5 h incubation of ^{15}N-labelled CsgB fibrils with D_2O, approximately 70% of the assigned resonances were significantly reduced or disappeared from the spectrum (**Supplementary figure 2B**), and after 24 h incubation, more than 85% of the resonances were significantly reduced or absent from the spectrum (**Supplementary figure 2C**). Significant reduction means that the peak height of each assigned resonance is at least less than 50% corresponding to its initial state (**Supplementary figure 2)**.

45% of CsgB residues displayed monoexponential exchange behavior (**Supplementary figure 3&4**), for example residues 72Ala, 94Asp and 110Gly (**Figure 36A**). 55% residues displayed biphasic exchange behavior (**Supplementary figure 3&4)**, for example, residue 42Ser, 59Gln and 148Val (**Figure 36C**). Most importantly, the H/D exchange behavior of selected residues of 42Ser and 59Gln, 72Ala, and 94Asp was similar with Nagaraj's processed data (**Figure 36;** Madhu Nagaraj, PhD thesis). It meant that two H/D data processing methods are identical.

Approximately 65% of the amyloid core residues were shown biphasic behavior. The fast exchange rates of these residues were in the range of 0.3 h^{-1} -30 h^{-1} and the slow rates were in the range of 1 h^{-1} to 0.001 h^{-1} (**Supplementary figure 3**).

In summary, the H/D exchange analysis of ^{15}N-labelled CsgB fibrils revealed that CsgB consists of an N-terminal (23Gly-34Ala) highly flexible region, an N-terminus single β-strand (35Val-41Ser), as well as an amyloid core. The amyloid core is composed of five repeating units (R1-R5), and each repeat contains two β-strands that are connected by a short β-turn. R5 is considered being involved in both binding to bacterial membrane and interacting with CsgF because of its high positively amide acid content (Madhu Nagaraj, PhD thesis).

5.3.4. H/D exchange NMR analysis of CsgB in mixed CsgA-^{15N}CsgB fibrils

The mixed CsgA-^{15N}CsgB fibrils were processed in the same manner as ^{15}N-labelled CsgB fibrils. In order to investigate the folding behavior of CsgB in mixed CsgA-CsgB fibrils, ^{15}N-labelled mixed fibrils were incubated in D_2O buffer for up to 6 w (0 h, 0.5 h, 1.5 h, 8 h, 24 h, 1 w, 2 w and 6 w) and processed by H/D exchange NMR. The spectra were processed by TopSpin. The H/D data were further analyzed and cacluated by the softwares of ccpNMR and Matlab.

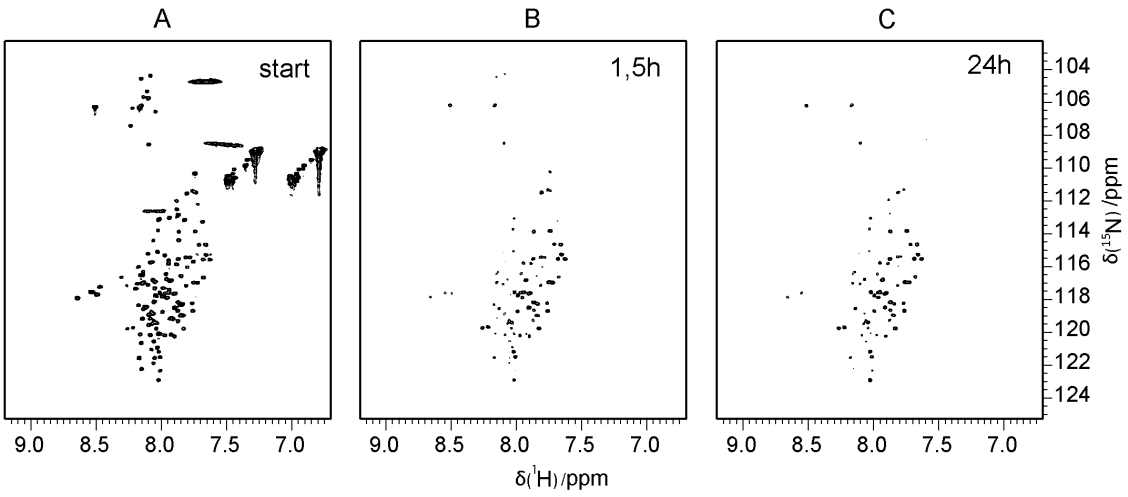


Figure 35: H/D exchange of mixed CsgA-^{15N}CsgB fibrils.

HSQC spectra of mixed CsgA-^{15N}CsgB fibrils (30 mg wet pellet of fibrils) dissolved in d6-DMSO containing 0.1 % (v/v) d1-TFA. **(A)** The spectra correspond to the initial protonated state (start) **(B)** 1.5 h exchanged fibrils and (**C**) 24 h exchanged fibrils.

The hydrogen exchange of the mixed CsgA-^{15N}CsgB fibrils resulted in HSQC spectra with different resonance intensities depending on the hydrogen bonding state of each backbone aminde, and incubation time in D_2O buffer. After 1.5 h incubation of mixed CsgA-^{15N}CsgB fibrils with D_2O, approximately 50% of the assigned CsgB resonances were significantly reduced or disappeared from the spectrum (**Figure 35B**), and after 24 h incubation, more than 65% of the resonances were significantly reduced or absent from the spectrum (**Figure 35C**). Significant reduction means that the peak height of each assigned resonance is at least less than 50% corresponding to its initial state (**Figure 35**).

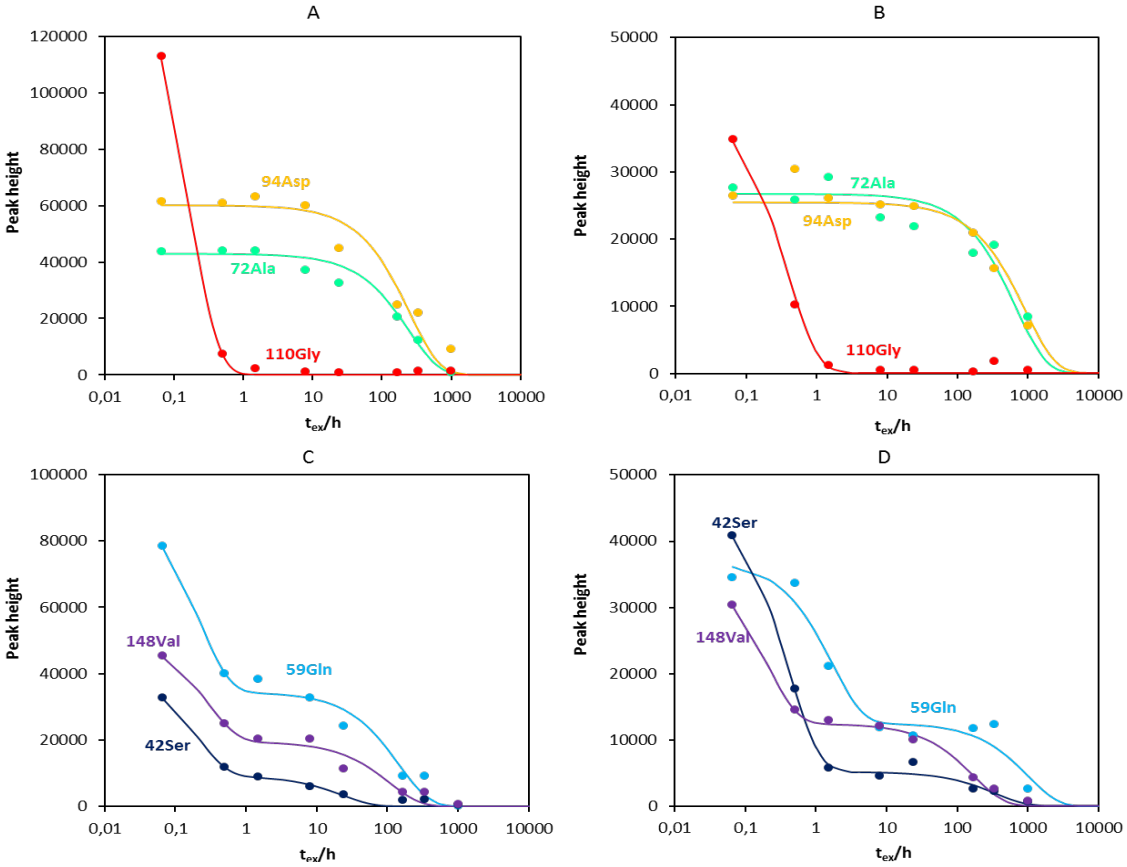


Figure 36: Determination of H/D exchange rates of selected CsgB residues in CsgB fibrils (A and C) and mixed CsgA-CsgB fibrils (B and D) respectively

Monoexponential fit and biphasic fit of peak heights at t_{DMSO}=0 h to t_{ex} corresponding to selected backbone amides of CsgB in CsgB fibrils (A and C) and mixed CsgA-CsgB firbils (B and D) respectively. The biphasic fit shows two different monoexponential equations.
Peak heights displaying corresponding to residues 72Ala, 94Asp and 110Gly showed monoexponential exchange behavior (A and B); Peak heights of residues 42Ser, 59Gln and 148Val showed biphasic exchange behavior (C and D).

All assigned CsgB resonances were evaluated and the corresponding H/D exchange rates were calculated. 60% of CsgB residues in mixed fibrils displayed a monoexponential dacay suggesting a homogenous structure of the fibrils (**Figure 37&38**), for example, residues of 72Ala, 94Asp and 110Gly (**Figure 36B**). The remaining of 40% CsgB residues in mixed fibrils displayed the biphasic behavior with a fast exchange rate and a slow exchange rate indicating heterogeneity of the fibrils (**Figure 37&38**), for example, residue 42Ser, 59Gln and 148Val (**Figure 36D**). The residue exhibiting biphasic behavior indicates presence of two different types of conformational structure (Yamaguchi, Katou et al. 2004, Luhrs, Ritter et al. 2005, Vilar, Chou et al. 2008).

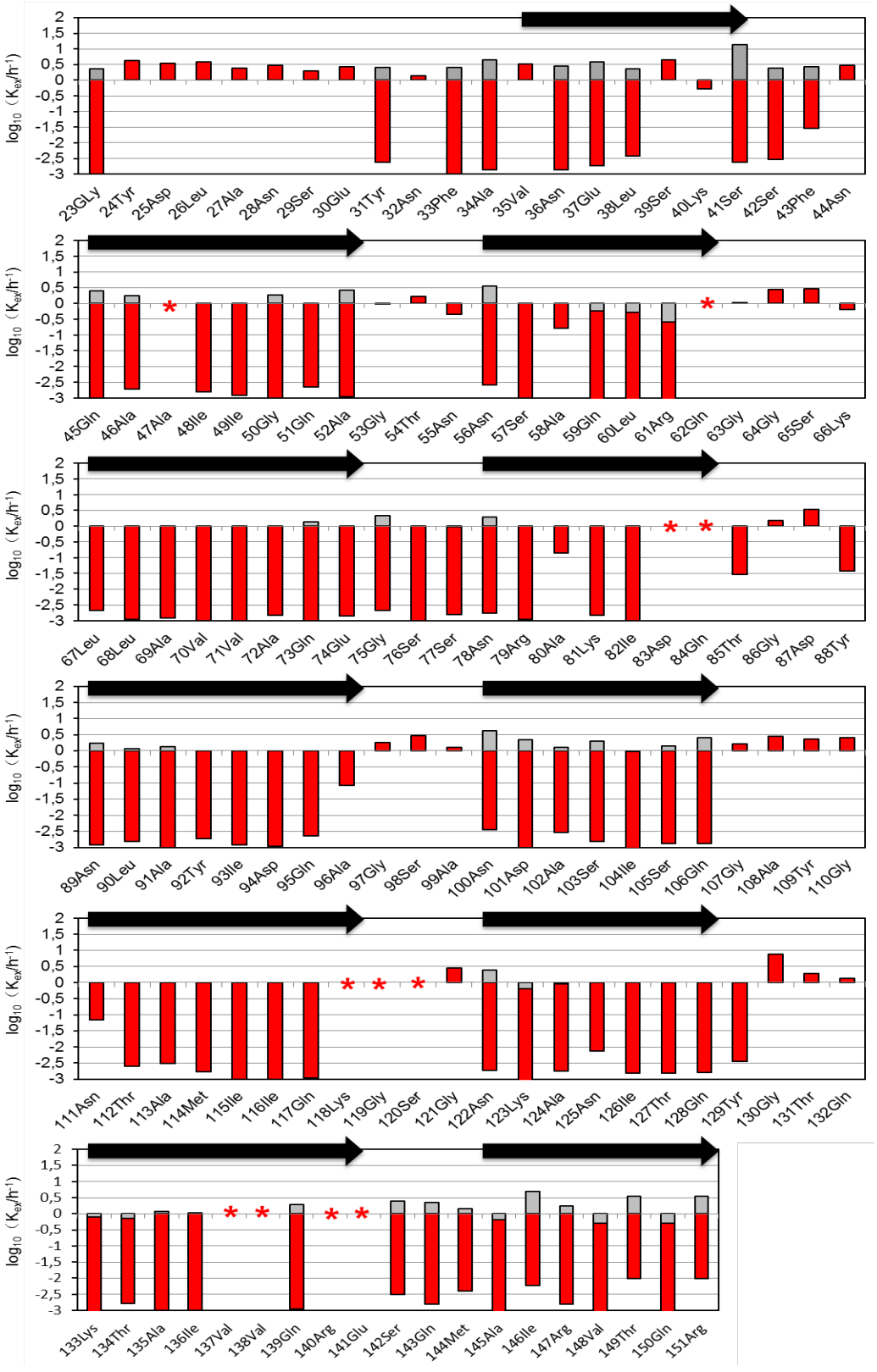


Figure 37: H/D exchange NMR data of CsgB in mixed CsgA-CsgB fibrils and secondary structure prediction

Red and grey bars indicated exchange rates (k_{ex} /h^{-1}) of CsgB. Gray bars indicated exchange rates of fast first phase in case of biphasic exchange behavior. Residues marked with a red star were missing assigments. Black arrows indicated β-strand.

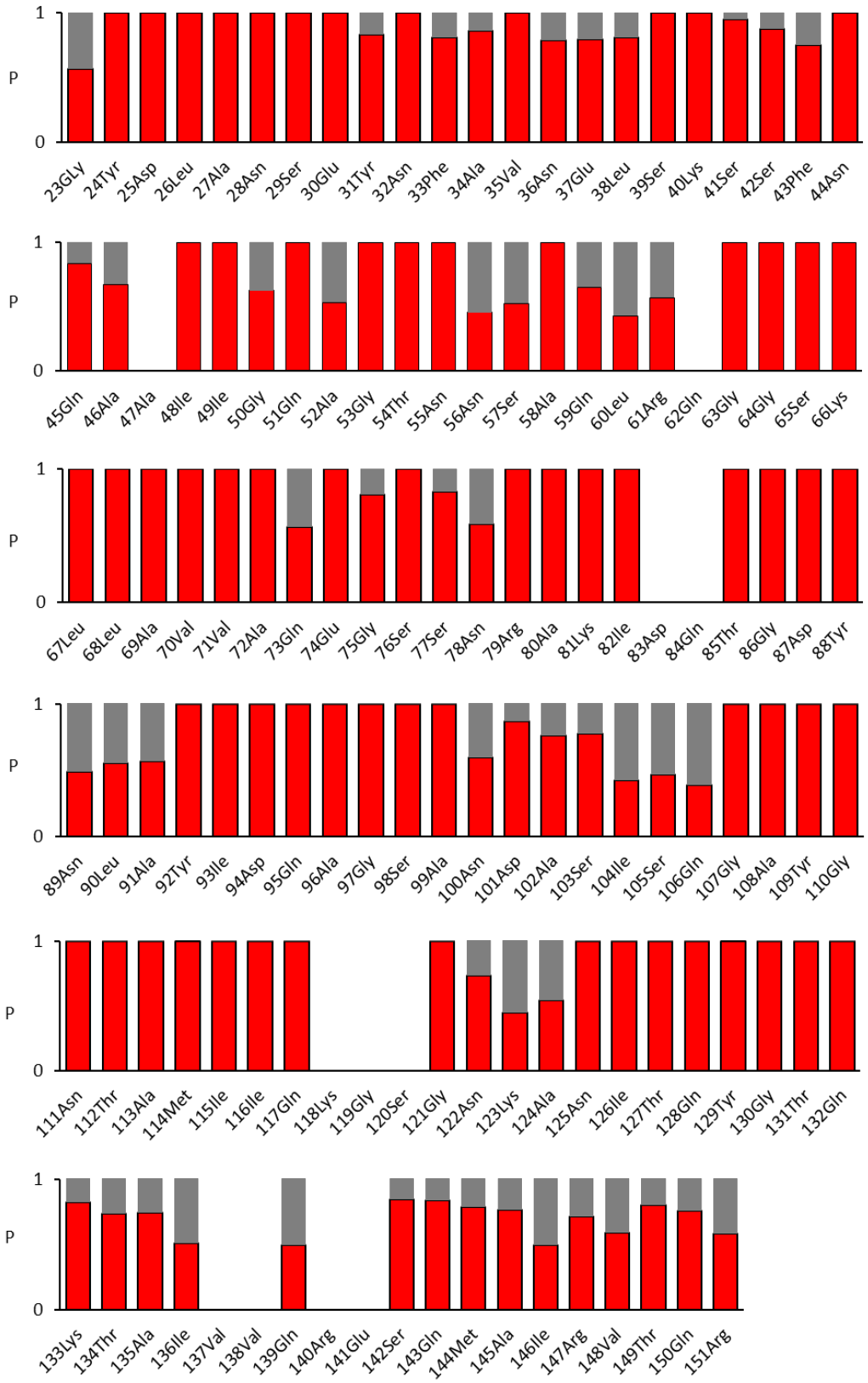


Figure 38: Relative populations in heterogenous fibers of CsgB in mixed fibrils

Grey bars indicate the minor populations (fast exchange) and green indicate major populations (slow exchange) in case of biphasic behavior.

H/D exchange rates and relative populations conresponding to all assigned residues exept 47Ala, 62Gln, 83Asp, 84Gln, 118Lys, 119Gly, 120Ser, 137Val, 138Val, 140Arg and 141Glu (because these residues cannot be assigned in the spectra) are shown in **Figure 37&38**. Red and grey bars indicated exchange rates of CsgB in mixed CsgA-CsgB fibrils. Gray bars indicated the fast exchange rates in the case of biphasic behavior. Amino-acid residues such as 44Asn, 49Ile, 58Ala, 70Ala, 82Ile and 117Gln showed monoexponential decay with exchange rates of 0.1 h^{-1} to 0.001 h^{-1}, indicating these amyloid core residues were identified to form a well-ordered and homogenous structure. However, relative majority residues belonging to the amyloid core such as 50Gly, 60Leu, 89Asn, 106Gln, 122Asn and 149Thr showed biphasic behavior indicating heterogeneity. The fast exchange rates of these residues were in the range of 0.3 h^{-1} -10 h^{-1} and the slow rates were in the range of 0.1 h^{-1} to 0.001 h^{-1}.
Residues belonging to either N-terminal flexible region or loop regions showed monoexponential decay with fast exchange rates (greater than 1 h^{-1}), for example, residues 24Tyr, 44Asn, 64Gly, 86Gly, 109Tyr and 121Gly. Some of the residues like Gly (63, 107 and 119), Thr (54, 85, 131 and 134), 30Glu, 38Leu, 43Phe, and 87Asp displayed a strong overlapping (**Supplementary figure 1**), for this reason, the accurate exchange rates of these residues cannot be calculated.
In summary, the H/D exchange analysis of mixed CsgA-^{15N}CsgB fibrils identified that CsgB consists of one of N-terminal single β-strand (35Val-41Ser) and five repeating units (R1-R5), each repeat is composed of two β-strands that are connected by a short β-turn. The residues belonging to β-strands were assumed to be involed in H-bonds that protect the corresponding backbone amids against proton exchange in D_2O buffer. Furthermore, the exchange data indicated that the N-terminus (23Gly-34Ala) is highly flexible.

5.3.5. In Comparison of the slow exchange rates between CsgB fibrils and mixed fibrils

In comparison with the slow exchange rates of amyloid core's residues of CsgB in pure CsgB fibrils, the most of the slow exchange rates in mixed firbils were significantly decreased (55 resdues/66 residues), this meant that CsgB in mixed fibrils formed a more ordered β-sheet structure (**Figure 39**), because amide-protons can well protect in the highly ordered β-sheet structure.

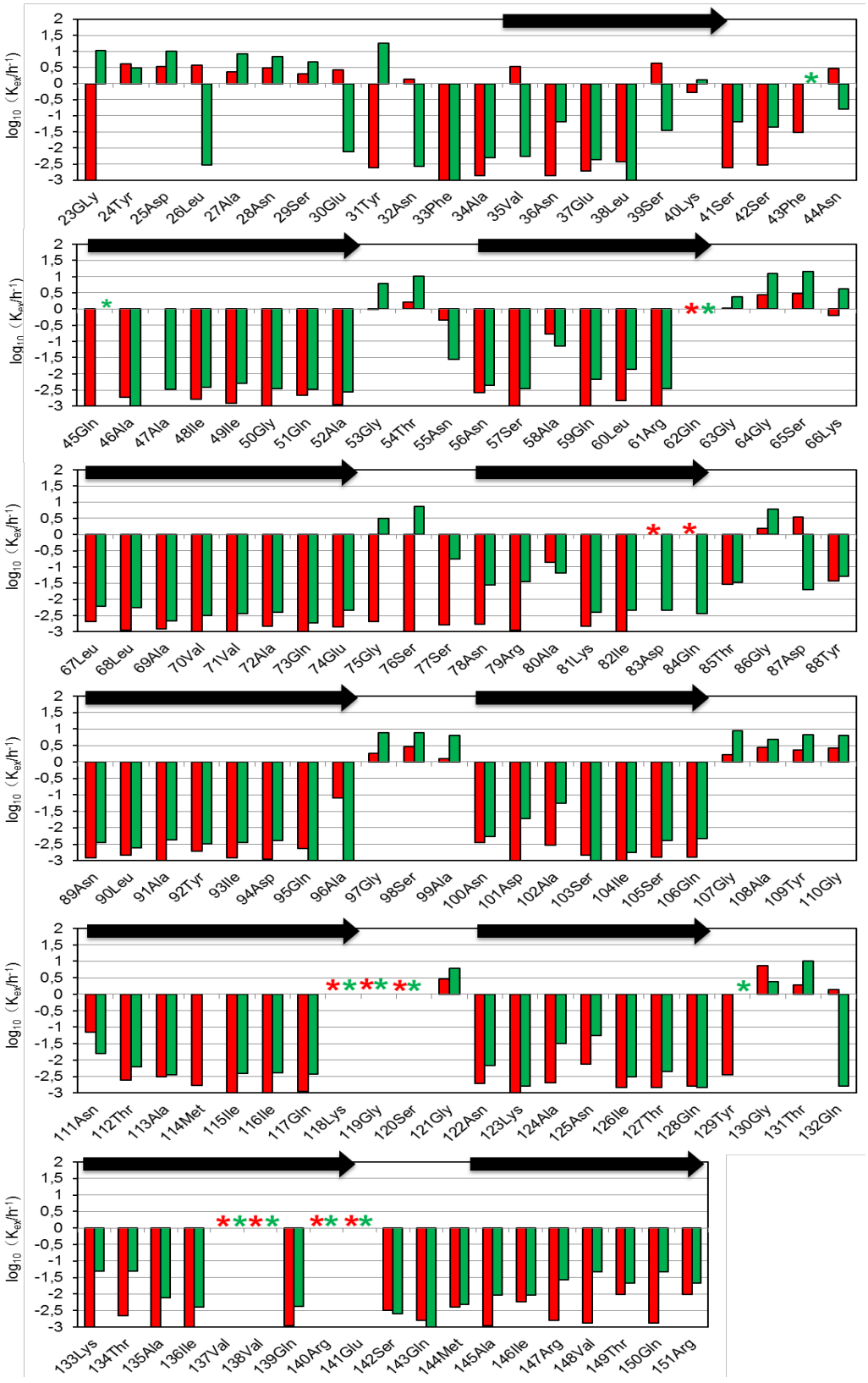


Figure 39: H/D exchange NMR data of CsgB and secondary structure prediction

Green bars indicated exchange rates (k_{ex}/ h^{-1}) of CsgB in CsgB fibrils and red bars indicated exchange rates (k_{ex}/ h^{-1}) of CsgB in mixed CsgA-CsgB fibrils. Residues marked with a green or red star were missing assigments. Black arrows indicated β-strands.

5.4. Disscussion

5.4.1. CsgA fibrils have the similar amyloidogenic properites formed in neutral and higher buffers

The ability of the major Curli subunit CsgA monomers to aggregate into amyloid-like fibrils was evaluated under a wide range of pH and salts conditions, CsgA displayed an inherent tendency to fibrillate (Wang, Hammer et al. 2008, Dueholm, Nielsen et al. 2011). In this thesis, CsgA fibrils formed in pH 7.4 buffer showed large clumps, but CsgA fibrils formed in pH 9 buffer showed no or less clumps (**Figure 22**). And the lag phase (10 μM CsgA) increased approimately 8-fold (**Figure 23**). These results indicated that the buffer pH could be a parameter for controlling for the aggregation of amyloid proteins. And CsgA fibrils formed in these two different buffers and exhibited the similar secondary structure (**Figure 24**).

5.4.2. Aggregation of CsgA protein can be effected by both CsgB protein and CsgA fibrils

Freshly buffer-exchanged CsgB protein and CsgA protein can co-fibrillize to form artificial Curli fibrils with any mixing ratio. In the co-aggregation experiments, the lag phase of aggregation was dramatically shortened in a concentration dependent manner of the CsgB (**Figure 25**). The mixed fibrils revealed the presence of amyloid-like fibers similar to CsgA fibrils and CsgB (**Figure 26**).

CsgA protein can be seeded by CsgA fibrils (2-day old) in pH 9.0 buffer, and the seeding capitcities depended on the concentration of seeds (1-8%) (**Figure 27**).

5.4.3. CsgB is located not only at the ends but also in between CsgA molecules in mixed fibrils

Amyloid fibrils can be detected by TEM with a high resoulution, but a single protein molecule is too small to be distinguished. Immunogold labelling TEM has been successfully used to study biochemical properties of amyloid-like fibrils such as Aβ (Romero, Aguilar et al. 2010, Al-Hilaly, Williams et al. 2013). Specific mixed fibrils containing 20% of biotinylated CsgB protein were investigated by immunogold labelling TEM after treated with gold-labelled biotin antibody. The TEM images clearly showed that CsgB molecule is not only

located the ends but also in between CsgA molecules in mixed fibrils with a certain intervals (**Figure 34**), and the interval can accommodate more than two CsgA molecules (**Figure 34**).

5.4.4. Interspersed CsgB could weaken the overall stability of Curli fibrils

Four mixed fibrils in presence of different ratios of CsgB and CsgA fibrils were fragmented into short fibrils by Selective Shearing Amplification array using a series of frequencies and various cycles. These processed fibrils were further used to seed CsgA protein followed by ThT fluorescence, and the lag times were calculated. During the CsgA seeding, each CsgA seed filament can be considered as a single end, monomeric CsgA adds to this end while starts the formation of a new CsgA fibril.

The lag time of CsgA seeded by processed CsgA fibrils (760Hz and 6 cycles) was only 9.8 hours, however, the lag time of (600 Hz and 6 cycles) was 18 hours and similar to the lag time (unprocessed CsgA fibrils; 18 hours) (**Figure 28**), it was assumed that CsgA fibrils could be broken up at the frequency of 760Hz or at the higher frequency for 6 cycles resulting in an extra number of fibril ends. For processed mixed CsgA-CsgB fibrils (containing 10% or 20% of CsgB) (**Figure 30&31**), interestingly, they could be broken up at the frequency of 600Hz or at the higher frequencies for 6 cycles. Taken together, 760Hz could be a critical frequency that can be used to differentiate pure CsgA fibrils and mixed fibrils.

In comparison to H/D exchange data of pure CsgA fibrils (Agnes Zimmer, PhD thesis) and pure CsgB fibrils (Madhu Nagaraj, PhD thesis), the slow H/D exchange rate of CsgA residues are much slower than the rate of CsgB residues, indicating the higher stability of CsgA fibrils than CsgB fibrils. These results might support that pure CsgB fibrils might be broken up using the critical frequency or lower frequencies.

In summary, mixed CsgA-CsgB fibrils appeared to be fragmented into optimal seeds at slightly lower shear forces compared to pure CsgA fibrils, possibly indicating that interspersed CsgB could weaken the overall stability of Curli fibrils.

5.4.5. CsgB in the mixed fibrils formed a more ordered β-sheet structure

H/D exchange NMR analysis revealed CsgB fibrils showing a high-level of heterogeneity. N-terminal flexible region (21-44) showed a single β-strand and R5 unit showed two β-strands with high level of heterogeneity, combined with the H/D exchange NMR analysis of ΔN & ΔR5 CsgB-truncation mutants, these results revealed that this specific structure of CsgB leads

to the skewed and overlapping arrangement of the β-strand, the N-terminus extra β-strand flips offer an interface for R5 unit of CsgA during the nucleation. And this arrangement helps not only in the CsgB nucleation of CsgA but also fibrils growth (Madhu Nagaraj, PhD thesis).

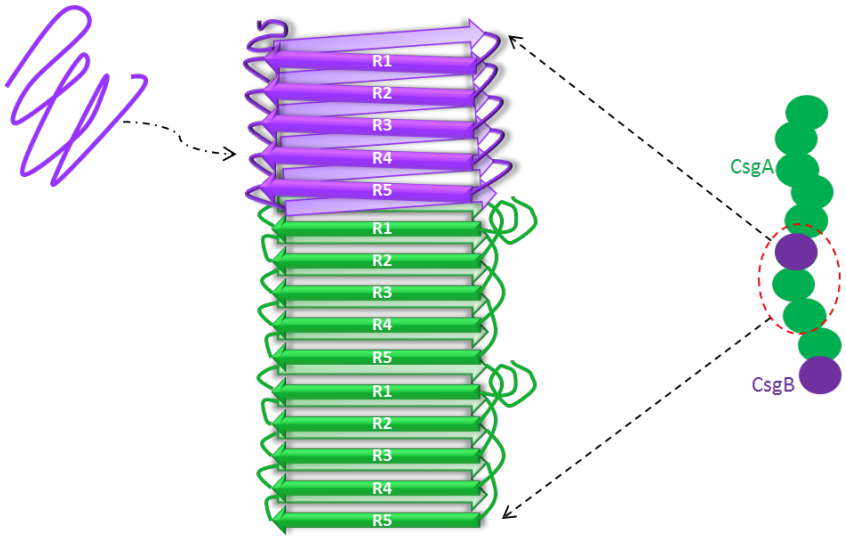


Figure 40: A suggested model for CsgA-CsgB interaction

A suggested model that CsgB is not only located the ends but also in between CsgA molecules in the fibrils. The CsgA R1 unit could provide a negatively charge interface for CsgB R5 unit. The combination between CsgA R1 and CsgB R5 may rely on charge interaction.

In comparison of H/D exchange data of CsgB fibrils and H/D exchange data of mixed fibrils (80% CsgA:20% CsgB), the H/D exchange rates of CsgB in mixed fibrils was slower than the exchange rates CsgB in CsgB fibrils (**Figure 37, Supplementary figure 3 and Figure 39**). 40% CsgB residues in mixed fibrils displayed biphasic exchange behavior (**Figure 38**), however, 55% residues in CsgB fibrils (**Supplementary figure 4**). The possible reason could interpret that CsgB forms a more ordered β-sheet structure in mixed fibrils because of the interaction of CsgA. Interestingly, in spite of CsgB R5 unit both in mixed fibrils and in CsgB fibrils displayed heterogeneous behavior, mixed fibrils slowed lower exchange rates (**Figure 39**) indicating a more stable structural state, this result could also demonstate that CsgB molecules place between CsgA molecules in mixed fibrils.

The sequence of CsgA R1 unit shows two negatively charged residues (24Glu and 43Asp), and the sequence of CsgB R5 unit exhibits four positively charged residues (132Arg, 139Arg, 147Arg and 151Arg) (**Figure 2**). It could explain a new interaction mannar between CsgA and CsgB during Curli assembly, the R1 unit of CsgA offers a negatively charge interface for CsgB R5 unit, and the interaction could be depended on the charge interaction (**Figure 40**).

Taken together:

1: CsgB forms a more ordered β-sheet structure in mixed fibrils.

2: R5 unit of CsgB could interact with neighbouring CsgA molecules. R1 unit of CsgA could offer a negatively charged interface and R5 of CsgB could provide a positively changed interface. Therefore, the interaction of C-terminus CsgB and N-terminus CsgA could depend on the role of positive and negative charges (**Figure 40**).

6. Conclusion

In this thesis, the isotope-labelling minmal medium was optimizated and the procedure of Curli purification was modified from the Collinson SK's method. Western-blot analysis demonstrates that native Curli contains approimately 90-95% of major subunit CsgA and approximately 5-10% of minor subunit CsgB. EM, FT-IR and solid-state NMR analysis demonstrate that native Curli displayed long thin unbranched fibrils with a well-ordered β-sheet secondry structure.

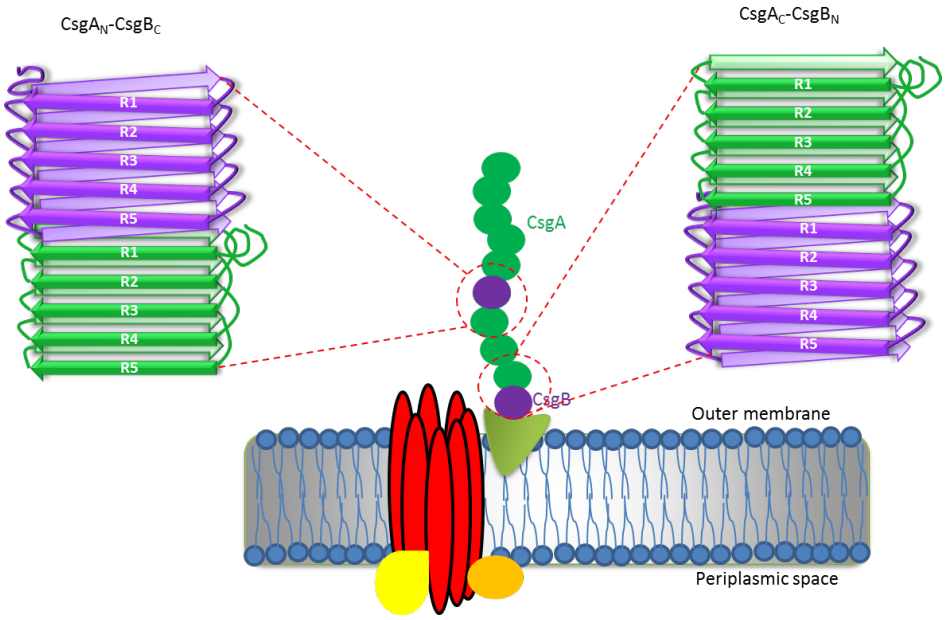


Figure 41: A suggested model for Curli assembly

The interaction between CsgA and CsgB has two different ways during Curli assembly. One way is that skewed and overlapping arrangement of N terminus single β-strand in CsgB provides a selective interface for interacting with C-terminus of CsgA, another way could be that C-terminus CsgB interacts with N-terminus CsgA though their charged interfaces.

In this work, Immungold labelling TEM images of mixed CsgA-CsgB fibrils clearly showed that CsgB is located not only at the ends but also in between CsgA molecules in mixed fibrils. Interspersed CsgB weakens the overall stability of Curli fibrils revealed by Selective Shearing Amplification array.

The H/D exchange data showed that fewer numbers of CsgB residues in mixed fibrils displayed biphasic exchange behavior compared to CsgB residues in pure CsgB fibrils, which could explain that interspersed CsgB formed a more ordered β-sheet structure. The results

could also support that C-terminus CsgB interacts with N-terminus CsgA through their charged interfaces (**Figure 41**). In addition, our colleages have found that skewed and overlapping arrangement of N-terminus single β-strand in CsgB provides a selective interface for interacting with C-terminus of CsgA (**Figure 41)** (Madhu Nagaraj, PhD thesis).

7. Future Directions

Isotope labelling native Curli was isolated from *E.coli*, solid-state NMR results of native Curli showed that it has the highly similarity of the β-sheet secondary structure with recombinantly CsgA fibrils. Furthermore, the standardized purification method of native Curli makes it possible to obtain a large amount of isotope labelled native Curli, and the H/D exchange NMR analysis of native Curli could reveal the sequence-specific structural information of Curli in the native manner.

We now know that CsgB is located not only at the ends but also in between CsgA molecules in mixed fibrils. In the future, the interaction of CsgA-CsgB can be investigated at atomic-level using native Curli or artificial mixed fibrils.

References

Al-Hilaly, Y. K., T. L. Williams, M. Stewart-Parker, L. Ford, E. Skaria, M. Cole, W. G. Bucher, K. L. Morris, A. A. Sada, J. R. Thorpe and L. C. Serpell (2013). "A central role for dityrosine crosslinking of Amyloid-beta in Alzheimer's disease." Acta Neuropathol Commun **1**(1): 83.

Anderson, G. G., J. J. Palermo, J. D. Schilling, R. Roth, J. Heuser and S. J. Hultgren (2003). "Intracellular bacterial biofilm-like pods in urinary tract infections." Science **301**(5629): 105-107.

Astbury, W. T., S. Dickinson and K. Bailey (1935). "The X-ray interpretation of denaturation and the structure of the seed globulins." Biochem J **29**(10): 2351-2360 2351.

Auluck, P. K., G. Caraveo and S. Lindquist (2010). "alpha-Synuclein: membrane interactions and toxicity in Parkinson's disease." Annu Rev Cell Dev Biol **26**: 211-233.

Bai, Y., J. S. Milne, L. Mayne and S. W. Englander (1993). "Primary structure effects on peptide group hydrogen exchange." Proteins **17**(1): 75-86.

Barnhart, M. M. and M. R. Chapman (2006). "Curli biogenesis and function." Annu Rev Microbiol **60**: 131-147.

Baxa, U., R. B. Wickner, A. C. Steven, D. E. Anderson, L. N. Marekov, W. M. Yau and R. Tycko (2007). "Characterization of beta-sheet structure in Ure2p1-89 yeast prion fibrils by solid-state nuclear magnetic resonance." Biochemistry **46**(45): 13149-13162.

Ben Nasr, A., A. Olsen, U. Sjobring, W. Muller-Esterl and L. Bjorck (1996). "Assembly of human contact phase proteins and release of bradykinin at the surface of curli-expressing Escherichia coli." Mol Microbiol **20**(5): 927-935.

Bian, Z., A. Brauner, Y. Li and S. Normark (2000). "Expression of and cytokine activation by Escherichia coli curli fibers in human sepsis." J Infect Dis **181**(2): 602-612.

Bian, Z. and S. Normark (1997). "Nucleator function of CsgB for the assembly of adhesive surface organelles in Escherichia coli." EMBO J **16**(19): 5827-5836.

Biancalana, M. and S. Koide (2010). "Molecular mechanism of Thioflavin-T binding to amyloid fibrils." Biochim Biophys Acta **1804**(7): 1405-1412.

Bjorkman, P. J., M. A. Saper, B. Samraoui, W. S. Bennett, J. L. Strominger and D. C. Wiley (1987). "Structure of the human class I histocompatibility antigen, HLA-A2." Nature **329**(6139): 506-512.

Brombacher, E., A. Baratto, C. Dorel and P. Landini (2006). "Gene expression regulation by the Curli activator CsgD protein: modulation of cellulose biosynthesis and control of negative determinants for microbial adhesion." J Bacteriol **188**(6): 2027-2037.

Castellino, F. J. and V. A. Ploplis (2005). "Structure and function of the plasminogen/plasmin system." Thromb Haemost **93**(4): 647-654.

Cegelski, L., J. S. Pinkner, N. D. Hammer, C. K. Cusumano, C. S. Hung, E. Chorell, V. Aberg, J. N. Walker, P. C. Seed, F. Almqvist, M. R. Chapman and S. J. Hultgren (2009). "Small-molecule inhibitors target Escherichia coli amyloid biogenesis and biofilm formation." Nat Chem Biol **5**(12): 913-919.

Chapman, M. R., L. S. Robinson, J. S. Pinkner, R. Roth, J. Heuser, M. Hammar, S. Normark and S. J. Hultgren (2002). "Role of Escherichia coli curli operons in directing amyloid fiber formation." Science **295**(5556): 851-855.

Chirwa, N. T. and M. B. Herrington (2003). "CsgD, a regulator of curli and cellulose synthesis, also regulates serine hydroxymethyltransferase synthesis in Escherichia coli K-12." Microbiology **149**(Pt 2): 525-535.

Chiti, F. and C. M. Dobson (2006). "Protein misfolding, functional amyloid, and human disease." Annu Rev Biochem **75**: 333-366.

Chromek, M., Z. Slamova, P. Bergman, L. Kovacs, L. Podracka, I. Ehren, T. Hokfelt, G. H.

Gudmundsson, R. L. Gallo, B. Agerberth and A. Brauner (2006). "The antimicrobial peptide cathelicidin protects the urinary tract against invasive bacterial infection." Nat Med **12**(6): 636-641.
Cohen, A. S. and E. Calkins (1959). "Electron microscopic observations on a fibrous component in amyloid of diverse origins." Nature **183**(4669): 1202-1203.
Collinson, S. K., P. C. Doig, J. L. Doran, S. Clouthier, T. J. Trust and W. W. Kay (1993). "Thin, aggregative fimbriae mediate binding of Salmonella enteritidis to fibronectin." J Bacteriol **175**(1): 12-18.
Collinson, S. K., L. Emody, K. H. Muller, T. J. Trust and W. W. Kay (1991). "Purification and characterization of thin, aggregative fimbriae from Salmonella enteritidis." J Bacteriol **173**(15): 4773-4781.
Collinson, S. K., J. M. Parker, R. S. Hodges and W. W. Kay (1999). "Structural predictions of AgfA, the insoluble fimbrial subunit of Salmonella thin aggregative fimbriae." J Mol Biol **290**(3): 741-756.
Czerski, L., O. Vinogradova and C. R. Sanders (2000). "NMR-Based amide hydrogen-deuterium exchange measurements for complex membrane proteins: development and critical evaluation." J Magn Reson **142**(1): 111-119.
Da Re, S. and J. M. Ghigo (2006). "A CsgD-independent pathway for cellulose production and biofilm formation in Escherichia coli." J Bacteriol **188**(8): 3073-3087.
Das, M., X. Mei, S. Jayaraman, D. Atkinson and O. Gursky (2014). "Amyloidogenic Mutations in Human Apolipoprotein A-I are not Necessarily Destabilizing: A Common Mechanism of ApoA-I Misfolding in Familial Amyloidosis and Atherosclerosis." FEBS J.
de Tullio, M. B., V. Castelletto, I. W. Hamley, P. V. Martino Adami, L. Morelli and E. M. Castano (2013). "Proteolytically inactive insulin-degrading enzyme inhibits amyloid formation yielding non-neurotoxic abeta peptide aggregates." PLoS One **8**(4): e59113.
Dueholm, M. S., S. B. Nielsen, K. L. Hein, P. Nissen, M. Chapman, G. Christiansen, P. H. Nielsen and D. E. Otzen (2011). "Fibrillation of the major curli subunit CsgA under a wide range of conditions implies a robust design of aggregation." Biochemistry **50**(39): 8281-8290.
Dzwolak, W., A. Loksztejn and V. Smirnovas (2006). "New insights into the self-assembly of insulin amyloid fibrils: an H-D exchange FT-IR study." Biochemistry **45**(26): 8143-8151.
Eckman, E. A., M. Watson, L. Marlow, K. Sambamurti and C. B. Eckman (2003). "Alzheimer's disease beta-amyloid peptide is increased in mice deficient in endothelin-converting enzyme." J Biol Chem **278**(4): 2081-2084.
Englander, S. W. (2000). "Protein folding intermediates and pathways studied by hydrogen exchange." Annu Rev Biophys Biomol Struct **29**: 213-238.
Epstein, E. A., M. A. Reizian and M. R. Chapman (2009). "Spatial clustering of the curlin secretion lipoprotein requires curli fiber assembly." J Bacteriol **191**(2): 608-615.
Fowler, D. M., A. V. Koulov, C. Alory-Jost, M. S. Marks, W. E. Balch and J. W. Kelly (2006). "Functional amyloid formation within mammalian tissue." PLoS Biol **4**(1): e6.
Fowler, D. M., A. V. Koulov, W. E. Balch and J. W. Kelly (2007). "Functional amyloid--from bacteria to humans." Trends Biochem Sci **32**(5): 217-224.
Fritzsching, K. J., Y. Yang, K. Schmidt-Rohr and M. Hong (2013). "Practical use of chemical shift databases for protein solid-state NMR: 2D chemical shift maps and amino-acid assignment with secondary-structure information." J Biomol NMR **56**(2): 155-167.
Gardiennet, C., A. K. Schutz, A. Hunkeler, B. Kunert, L. Terradot, A. Bockmann and B. H. Meier (2012). "A sedimented sample of a 59 kDa dodecameric helicase yields high-resolution solid-state NMR spectra." Angew Chem Int Ed Engl **51**(31): 7855-7858.
Gebbink, M. F., D. Claessen, B. Bouma, L. Dijkhuizen and H. A. Wosten (2005). "Amyloids--a functional coat for microorganisms." Nat Rev Microbiol **3**(4): 333-341.
Germain, R. N. and D. H. Margulies (1993). "The biochemistry and cell biology of antigen processing and presentation." Annu Rev Immunol **11**: 403-450.

Gerstel, U., C. Park and U. Romling (2003). "Complex regulation of csgD promoter activity by global regulatory proteins." Mol Microbiol **49**(3): 639-654.
Giliano, N. Y., S. I. Stepanov, L. A. Noskin, E. N. Arkhipova and L. V. Konevega (2013). "[Assessing of the role superoxide, nitric oxide and redox metals in cytotoxic effect of the H2O2 and amyloid-beta-protein]." Patol Fiziol Eksp Ter(3): 42-46.
Gophna, U., T. A. Oelschlaeger, J. Hacker and E. Z. Ron (2002). "Role of fibronectin in curli-mediated internalization." FEMS Microbiol Lett **212**(1): 55-58.
Goyal, P., P. V. Krasteva, N. Van Gerven, F. Gubellini, I. Van den Broeck, A. Troupiotis-Tsailaki, W. Jonckheere, G. Pehau-Arnaudet, J. S. Pinkner, M. R. Chapman, S. J. Hultgren, S. Howorka, R. Fronzes and H. Remaut (2014). "Structural and mechanistic insights into the bacterial amyloid secretion channel CsgG." Nature **516**(7530): 250-253.
Greenwald, J. and R. Riek (2010). "Biology of amyloid: structure, function, and regulation." Structure **18**(10): 1244-1260.
Hammar, M., Z. Bian and S. Normark (1996). "Nucleator-dependent intercellular assembly of adhesive curli organelles in Escherichia coli." Proc Natl Acad Sci U S A **93**(13): 6562-6566.
Hammer, N. D., B. A. McGuffie, Y. Zhou, M. P. Badtke, A. A. Reinke, K. Brannstrom, J. E. Gestwicki, A. Olofsson, F. Almqvist and M. R. Chapman (2012). "The C-terminal repeating units of CsgB direct bacterial functional amyloid nucleation." J Mol Biol **422**(3): 376-389.
Hammer, N. D., J. C. Schmidt and M. R. Chapman (2007). "The curli nucleator protein, CsgB, contains an amyloidogenic domain that directs CsgA polymerization." Proc Natl Acad Sci U S A **104**(30): 12494-12499.
Harrison, R. S., P. C. Sharpe, Y. Singh and D. P. Fairlie (2007). "Amyloid peptides and proteins in review." Rev Physiol Biochem Pharmacol **159**: 1-77.
Hazenberg, B. P. (2013). "Amyloidosis: a clinical overview." Rheum Dis Clin North Am **39**(2): 323-345.
Hensley, K., J. M. Carney, M. P. Mattson, M. Aksenova, M. Harris, J. F. Wu, R. A. Floyd and D. A. Butterfield (1994). "A model for beta-amyloid aggregation and neurotoxicity based on free radical generation by the peptide: relevance to Alzheimer disease." Proc Natl Acad Sci U S A **91**(8): 3270-3274.
Herczenik, E. and M. F. Gebbink (2008). "Molecular and cellular aspects of protein misfolding and disease." FASEB J **22**(7): 2115-2133.
Herwald, H., M. Morgelin, A. Olsen, M. Rhen, B. Dahlback, W. Muller-Esterl and L. Bjorck (1998). "Activation of the contact-phase system on bacterial surfaces--a clue to serious complications in infectious diseases." Nat Med **4**(3): 298-302.
Hiramatsu, H. and T. Kitagawa (2005). "FT-IR approaches on amyloid fibril structure." Biochim Biophys Acta **1753**(1): 100-107.
Holmqvist, E., J. Reimegard, M. Sterk, N. Grantcharova, U. Romling and E. G. Wagner (2010). "Two antisense RNAs target the transcriptional regulator CsgD to inhibit curli synthesis." EMBO J **29**(11): 1840-1850.
Hoshino, M., H. Katou, Y. Hagihara, K. Hasegawa, H. Naiki and Y. Goto (2002). "Mapping the core of the beta(2)-microglobulin amyloid fibril by H/D exchange." Nat Struct Biol **9**(5): 332-336.
Howie, A. J. and D. B. Brewer (2009). "Optical properties of amyloid stained by Congo red: history and mechanisms." Micron **40**(3): 285-301.
Huber, M., O. Y. Ovchinnikova, A. K. Schutz, R. Glockshuber, B. H. Meier and A. Bockmann (2014). "Solid-state NMR sequential assignment of Osaka-mutant amyloid-beta (Abeta1-40 E22Delta) fibrils." Biomol NMR Assign.
Hung, C., J. Marschall, C. A. Burnham, A. S. Byun and J. P. Henderson (2014). "The bacterial amyloid curli is associated with urinary source bloodstream infection." PLoS One **9**(1): e86009.
Ippel, J. H., A. Olofsson, J. Schleucher, E. Lundgren and S. S. Wijmenga (2002). "Probing

solvent accessibility of amyloid fibrils by solution NMR spectroscopy." Proc Natl Acad Sci U S A **99**(13): 8648-8653.
Iwata, K., T. Fujiwara, Y. Matsuki, H. Akutsu, S. Takahashi, H. Naiki and Y. Goto (2006). "3D structure of amyloid protofilaments of beta2-microglobulin fragment probed by solid-state NMR." Proc Natl Acad Sci U S A **103**(48): 18119-18124.
Jimenez, J. L., J. I. Guijarro, E. Orlova, J. Zurdo, C. M. Dobson, M. Sunde and H. R. Saibil (1999). "Cryo-electron microscopy structure of an SH3 amyloid fibril and model of the molecular packing." EMBO J **18**(4): 815-821.
Jimenez, J. L., E. J. Nettleton, M. Bouchard, C. V. Robinson, C. M. Dobson and H. R. Saibil (2002). "The protofilament structure of insulin amyloid fibrils." Proc Natl Acad Sci U S A **99**(14): 9196-9201.
Johansson, C., T. Nilsson, A. Olsen and M. J. Wick (2001). "The influence of curli, a MHC-I-binding bacterial surface structure, on macrophage-T cell interactions." FEMS Immunol Med Microbiol **30**(1): 21-29.
Jorgensen, T. J., N. Bache, P. Roepstorff, H. Gardsvoll and M. Ploug (2005). "Collisional activation by MALDI tandem time-of-flight mass spectrometry induces intramolecular migration of amide hydrogens in protonated peptides." Mol Cell Proteomics **4**(12): 1910-1919.
Jorgensen, T. J., H. Gardsvoll, M. Ploug and P. Roepstorff (2005). "Intramolecular migration of amide hydrogens in protonated peptides upon collisional activation." J Am Chem Soc **127**(8): 2785-2793.
Jubelin, G., A. Vianney, C. Beloin, J. M. Ghigo, J. C. Lazzaroni, P. Lejeune and C. Dorel (2005). "CpxR/OmpR interplay regulates curli gene expression in response to osmolarity in Escherichia coli." J Bacteriol **187**(6): 2038-2049.
Justice, S. S., C. Hung, J. A. Theriot, D. A. Fletcher, G. G. Anderson, M. J. Footer and S. J. Hultgren (2004). "Differentiation and developmental pathways of uropathogenic Escherichia coli in urinary tract pathogenesis." Proc Natl Acad Sci U S A **101**(5): 1333-1338.
Kai-Larsen, Y., P. Luthje, M. Chromek, V. Peters, X. Wang, A. Holm, L. Kadas, K. O. Hedlund, J. Johansson, M. R. Chapman, S. H. Jacobson, U. Romling, B. Agerberth and A. Brauner (2010). "Uropathogenic Escherichia coli modulates immune responses and its curli fimbriae interact with the antimicrobial peptide LL-37." PLoS Pathog **6**(7): e1001010.
Khurana, R., V. N. Uversky, L. Nielsen and A. L. Fink (2001). "Is Congo red an amyloid-specific dye?" J Biol Chem **276**(25): 22715-22721.
Kolezhuk, A. K. (1996). "Continuum-field description of one-dimensional dimerized spin-1/2 Heisenberg antiferromagnets." Phys Rev B Condens Matter **53**(1): 318-327.
Kong, J. and S. Yu (2007). "Fourier transform infrared spectroscopic analysis of protein secondary structures." Acta Biochim Biophys Sin (Shanghai) **39**(8): 549-559.
Kyle, R. A. (2001). "Amyloidosis: a convoluted story." Br J Haematol **114**(3): 529-538.
Lange, A., Z. Gattin, H. Van Melckebeke, C. Wasmer, A. Soragni, W. F. van Gunsteren and B. H. Meier (2009). "A combined solid-state NMR and MD characterization of the stability and dynamics of the HET-s(218-289) prion in its amyloid conformation." Chembiochem **10**(10): 1657-1665.
Li, J. J., G. Dolios, R. Wang and F. F. Liao (2014). "Soluble beta-amyloid peptides, but not insoluble fibrils, have specific effect on neuronal microRNA expression." PLoS One **9**(3): e90770.
Lim, K. H., T. N. Nguyen, S. M. Damo, T. Mazur, H. L. Ball, S. B. Prusiner, A. Pines and D. E. Wemmer (2006). "Solid-state NMR structural studies of the fibril form of a mutant mouse prion peptide PrP89-143(P101L)." Solid State Nucl Magn Reson **29**(1-3): 183-190.
Luhrs, T., C. Ritter, M. Adrian, D. Riek-Loher, B. Bohrmann, H. Dobeli, D. Schubert and R. Riek (2005). "3D structure of Alzheimer's amyloid-beta(1-42) fibrils." Proc Natl Acad Sci U S A **102**(48): 17342-17347.
Macarisin, D., J. Patel, G. Bauchan, J. A. Giron and V. K. Sharma (2012). "Role of curli and

cellulose expression in adherence of Escherichia coli O157:H7 to spinach leaves." Foodborne Pathog Dis **9**(2): 160-167.
Malinchik, S. B., H. Inouye, K. E. Szumowski and D. A. Kirschner (1998). "Structural analysis of Alzheimer's beta(1-40) amyloid: protofilament assembly of tubular fibrils." Biophys J **74**(1): 537-545.
Martin, S. E., M. D. Benson and E. M. Hattab (2014). "The pathologic spectrum of oculoleptomeningeal amyloidosis with Val30Gly transthyretin gene mutation in a postmortem case." Hum Pathol.
McCord, L. A., W. G. Liang, E. Dowdell, V. Kalas, R. J. Hoey, A. Koide, S. Koide and W. J. Tang (2013). "Conformational states and recognition of amyloidogenic peptides of human insulin-degrading enzyme." Proc Natl Acad Sci U S A **110**(34): 13827-13832.
Mollica, A., A. Stefanucci, R. Costante and F. Pinnen (2012). "Role of formyl peptide receptors (FPR) in abnormal inflammation responses involved in neurodegenerative diseases." Antiinflamm Antiallergy Agents Med Chem **11**(1): 20-36.
Nabuurs, S. M. and C. P. van Mierlo (2010). "Interrupted hydrogen/deuterium exchange reveals the stable core of the remarkably helical molten globule of alpha-beta parallel protein flavodoxin." J Biol Chem **285**(6): 4165-4172.
Nelson, R., M. R. Sawaya, M. Balbirnie, A. O. Madsen, C. Riekel, R. Grothe and D. Eisenberg (2005). "Structure of the cross-beta spine of amyloid-like fibrils." Nature **435**(7043): 773-778.
Nenninger, A. A., L. S. Robinson, N. D. Hammer, E. A. Epstein, M. P. Badtke, S. J. Hultgren and M. R. Chapman (2011). "CsgE is a curli secretion specificity factor that prevents amyloid fibre aggregation." Mol Microbiol **81**(2): 486-499.
Nenninger, A. A., L. S. Robinson and S. J. Hultgren (2009). "Localized and efficient curli nucleation requires the chaperone-like amyloid assembly protein CsgF." Proc Natl Acad Sci U S A **106**(3): 900-905.
Nielsen, J. T., M. Bjerring, M. D. Jeppesen, R. O. Pedersen, J. M. Pedersen, K. L. Hein, T. Vosegaard, T. Skrydstrup, D. E. Otzen and N. C. Nielsen (2009). "Unique identification of supramolecular structures in amyloid fibrils by solid-state NMR spectroscopy." Angew Chem Int Ed Engl **48**(12): 2118-2121.
Olofsson, A., J. H. Ippel, S. S. Wijmenga, E. Lundgren and A. Ohman (2004). "Probing solvent accessibility of transthyretin amyloid by solution NMR spectroscopy." J Biol Chem **279**(7): 5699-5707.
Olofsson, A., M. Lindhagen-Persson, M. Vestling, A. E. Sauer-Eriksson and A. Ohman (2009). "Quenched hydrogen/deuterium exchange NMR characterization of amyloid-beta peptide aggregates formed in the presence of Cu2+ or Zn2+." FEBS J **276**(15): 4051-4060.
Olsen, A., A. Arnqvist, M. Hammar and S. Normark (1993). "Environmental regulation of curli production in Escherichia coli." Infect Agents Dis **2**(4): 272-274.
Olsen, A., H. Herwald, M. Wikstrom, K. Persson, E. Mattsson and L. Bjorck (2002). "Identification of two protein-binding and functional regions of curli, a surface organelle and virulence determinant of Escherichia coli." J Biol Chem **277**(37): 34568-34572.
Olsen, A., A. Jonsson and S. Normark (1989). "Fibronectin binding mediated by a novel class of surface organelles on Escherichia coli." Nature **338**(6217): 652-655.
Olsen, A., M. J. Wick, M. Morgelin and L. Bjorck (1998). "Curli, fibrous surface proteins of Escherichia coli, interact with major histocompatibility complex class I molecules." Infect Immun **66**(3): 944-949.
Olubiyi, O. O., D. Frenzel, D. Bartnik, J. M. Gluck, O. Brener, L. Nagel-Steger, S. A. Funke, D. Willbold and B. Strodel (2013). "Amyloid Aggregation Inhibitory Mechanism of Arginine-Rich D-Peptides." Curr Med Chem.
Ono, K., M. M. Condron and D. B. Teplow (2009). "Structure-neurotoxicity relationships of amyloid beta-protein oligomers." Proc Natl Acad Sci U S A **106**(35): 14745-14750.

Otzen, D. and P. H. Nielsen (2008). "We find them here, we find them there: functional bacterial amyloid." Cell Mol Life Sci **65**(6): 910-927.
Palmer, J. C., S. Baig, P. G. Kehoe and S. Love (2009). "Endothelin-converting enzyme-2 is increased in Alzheimer's disease and up-regulated by Abeta." Am J Pathol **175**(1): 262-270.
Palmer, J. C., H. M. Tayler and S. Love (2013). "Endothelin-converting enzyme-1 activity, endothelin-1 production, and free radical-dependent vasoconstriction in Alzheimer's disease." J Alzheimers Dis **36**(3): 577-587.
Parthasarathy, S., F. Long, Y. Miller, Y. Xiao, D. McElheny, K. Thurber, B. Ma, R. Nussinov and Y. Ishii (2011). "Molecular-level examination of Cu2+ binding structure for amyloid fibrils of 40-residue Alzheimer's beta by solid-state NMR spectroscopy." J Am Chem Soc **133**(10): 3390-3400.
Pike, C. J., A. J. Walencewicz, C. G. Glabe and C. W. Cotman (1991). "In vitro aging of beta-amyloid protein causes peptide aggregation and neurotoxicity." Brain Res **563**(1-2): 311-314.
Rapsinski, G. J., T. N. Newman, G. O. Oppong, J. P. van Putten and C. Tukel (2013). "CD14 protein acts as an adaptor molecule for the immune recognition of Salmonella curli fibers." J Biol Chem **288**(20): 14178-14188.
Ritter, C., M. L. Maddelein, A. B. Siemer, T. Luhrs, M. Ernst, B. H. Meier, S. J. Saupe and R. Riek (2005). "Correlation of structural elements and infectivity of the HET-s prion." Nature **435**(7043): 844-848.
Robbins, K. C., L. Summaria, B. Hsieh and R. J. Shah (1967). "The peptide chains of human plasmin. Mechanism of activation of human plasminogen to plasmin." J Biol Chem **242**(10): 2333-2342.
Robinson, L. S., E. M. Ashman, S. J. Hultgren and M. R. Chapman (2006). "Secretion of curli fibre subunits is mediated by the outer membrane-localized CsgG protein." Mol Microbiol **59**(3): 870-881.
Romero, D., C. Aguilar, R. Losick and R. Kolter (2010). "Amyloid fibers provide structural integrity to Bacillus subtilis biofilms." Proc Natl Acad Sci U S A **107**(5): 2230-2234.
Romling, U., Z. Bian, M. Hammar, W. D. Sierralta and S. Normark (1998). "Curli fibers are highly conserved between Salmonella typhimurium and Escherichia coli with respect to operon structure and regulation." J Bacteriol **180**(3): 722-731.
Sabate, R., U. Baxa, L. Benkemoun, N. Sanchez de Groot, B. Coulary-Salin, M. L. Maddelein, L. Malato, S. Ventura, A. C. Steven and S. J. Saupe (2007). "Prion and non-prion amyloids of the HET-s prion forming domain." J Mol Biol **370**(4): 768-783.
Saksela, O. and D. B. Rifkin (1988). "Cell-associated plasminogen activation: regulation and physiological functions." Annu Rev Cell Biol **4**: 93-126.
Saldana, Z., J. Xicohtencatl-Cortes, F. Avelino, A. D. Phillips, J. B. Kaper, J. L. Puente and J. A. Giron (2009). "Synergistic role of curli and cellulose in cell adherence and biofilm formation of attaching and effacing Escherichia coli and identification of Fis as a negative regulator of curli." Environ Microbiol **11**(4): 992-1006.
Scheidt, H. A., I. Morgado and D. Huster (2012). "Solid-state NMR reveals a close structural relationship between amyloid-beta protofibrils and oligomers." J Biol Chem **287**(27): 22822-22826.
Schutz, A. K., A. Soragni, S. Hornemann, A. Aguzzi, M. Ernst, A. Bockmann and B. H. Meier (2011). "The amyloid-Congo red interface at atomic resolution." Angew Chem Int Ed Engl **50**(26): 5956-5960.
Shewmaker, F., R. P. McGlinchey, K. R. Thurber, P. McPhie, F. Dyda, R. Tycko and R. B. Wickner (2009). "The functional curli amyloid is not based on in-register parallel beta-sheet structure." J Biol Chem **284**(37): 25065-25076.
Sipe, J. D. and A. S. Cohen (2000). "Review: history of the amyloid fibril." J Struct Biol **130**(2-3): 88-98.
Sjobring, U., G. Pohl and A. Olsen (1994). "Plasminogen, absorbed by Escherichia coli

expressing curli or by Salmonella enteritidis expressing thin aggregative fimbriae, can be activated by simultaneously captured tissue-type plasminogen activator (t-PA)." Mol Microbiol **14**(3): 443-452.
Sultana, R., P. Mecocci, F. Mangialasche, R. Cecchetti, M. Baglioni and D. A. Butterfield (2011). "Increased protein and lipid oxidative damage in mitochondria isolated from lymphocytes from patients with Alzheimer's disease: insights into the role of oxidative stress in Alzheimer's disease and initial investigations into a potential biomarker for this dementing disorder." J Alzheimers Dis **24**(1): 77-84.
Sunde, M., L. C. Serpell, M. Bartlam, P. E. Fraser, M. B. Pepys and C. C. Blake (1997). "Common core structure of amyloid fibrils by synchrotron X-ray diffraction." J Mol Biol **273**(3): 729-739.
Takeuchi, O. and S. Akira (2010). "Pattern recognition receptors and inflammation." Cell **140**(6): 805-820.
Taylor, J. D., Y. Zhou, P. S. Salgado, A. Patwardhan, M. McGuffie, T. Pape, G. Grabe, E. Ashman, S. C. Constable, P. J. Simpson, W. C. Lee, E. Cota, M. R. Chapman and S. J. Matthews (2011). "Atomic resolution insights into curli fiber biogenesis." Structure **19**(9): 1307-1316.
Townsend, M., G. M. Shankar, T. Mehta, D. M. Walsh and D. J. Selkoe (2006). "Effects of secreted oligomers of amyloid beta-protein on hippocampal synaptic plasticity: a potent role for trimers." J Physiol **572**(Pt 2): 477-492.
Toyama, B. H. and J. S. Weissman (2011). "Amyloid structure: conformational diversity and consequences." Annu Rev Biochem **80**: 557-585.
Trautner, B. W. and R. O. Darouiche (2004). "Role of biofilm in catheter-associated urinary tract infection." Am J Infect Control **32**(3): 177-183.
Tukel, C., J. H. Nishimori, R. P. Wilson, M. G. Winter, A. M. Keestra, J. P. van Putten and A. J. Baumler (2010). "Toll-like receptors 1 and 2 cooperatively mediate immune responses to curli, a common amyloid from enterobacterial biofilms." Cell Microbiol **12**(10): 1495-1505.
Tukel, C., M. Raffatellu, A. D. Humphries, R. P. Wilson, H. L. Andrews-Polymenis, T. Gull, J. F. Figueiredo, M. H. Wong, K. S. Michelsen, M. Akcelik, L. G. Adams and A. J. Baumler (2005). "CsgA is a pathogen-associated molecular pattern of Salmonella enterica serotype Typhimurium that is recognized by Toll-like receptor 2." Mol Microbiol **58**(1): 289-304.
Tukel, C., R. P. Wilson, J. H. Nishimori, M. Pezeshki, B. A. Chromy and A. J. Baumler (2009). "Responses to amyloids of microbial and host origin are mediated through toll-like receptor 2." Cell Host Microbe **6**(1): 45-53.
Van Melckebeke, H., P. Schanda, J. Gath, C. Wasmer, R. Verel, A. Lange, B. H. Meier and A. Bockmann (2011). "Probing water accessibility in HET-s(218-289) amyloid fibrils by solid-state NMR." J Mol Biol **405**(3): 765-772.
Van Melckebeke, H., C. Wasmer, A. Lange, E. Ab, A. Loquet, A. Bockmann and B. H. Meier (2010). "Atomic-resolution three-dimensional structure of HET-s(218-289) amyloid fibrils by solid-state NMR spectroscopy." J Am Chem Soc **132**(39): 13765-13775.
Vassar, P. S. and C. F. Culling (1959). "Fluorescent stains, with special reference to amyloid and connective tissues." Arch Pathol **68**: 487-498.
Vilar, M., H. T. Chou, T. Luhrs, S. K. Maji, D. Riek-Loher, R. Verel, G. Manning, H. Stahlberg and R. Riek (2008). "The fold of alpha-synuclein fibrils." Proc Natl Acad Sci U S A **105**(25): 8637-8642.
Vilar, M., L. Wang and R. Riek (2012). "Structural studies of amyloids by quenched hydrogen-deuterium exchange by NMR." Methods Mol Biol **849**: 185-198.
Wang, X. and M. R. Chapman (2008). "Sequence determinants of bacterial amyloid formation." J Mol Biol **380**(3): 570-580.
Wang, X., N. D. Hammer and M. R. Chapman (2008). "The molecular basis of functional bacterial amyloid polymerization and nucleation." J Biol Chem **283**(31): 21530-21539.

Wang, X., D. R. Smith, J. W. Jones and M. R. Chapman (2007). "In vitro polymerization of a functional Escherichia coli amyloid protein." J Biol Chem **282**(6): 3713-3719.
Wang, X., Y. Zhou, J. J. Ren, N. D. Hammer and M. R. Chapman (2010). "Gatekeeper residues in the major curlin subunit modulate bacterial amyloid fiber biogenesis." Proc Natl Acad Sci U S A **107**(1): 163-168.
Wasmer, C., L. Benkemoun, R. Sabate, M. O. Steinmetz, B. Coulary-Salin, L. Wang, R. Riek, S. J. Saupe and B. H. Meier (2009). "Solid-state NMR spectroscopy reveals that E. coli inclusion bodies of HET-s(218-289) are amyloids." Angew Chem Int Ed Engl **48**(26): 4858-4860.
Wasmer, C., A. Lange, H. Van Melckebeke, A. B. Siemer, R. Riek and B. H. Meier (2008). "Amyloid fibrils of the HET-s(218-289) prion form a beta solenoid with a triangular hydrophobic core." Science **319**(5869): 1523-1526.
White, A. P., S. K. Collinson, P. A. Banser, D. L. Gibson, M. Paetzel, N. C. Strynadka and W. W. Kay (2001). "Structure and characterization of AgfB from Salmonella enteritidis thin aggregative fimbriae." J Mol Biol **311**(4): 735-749.
Wickner, R. B., F. Dyda and R. Tycko (2008). "Amyloid of Rnq1p, the basis of the [PIN+] prion, has a parallel in-register beta-sheet structure." Proc Natl Acad Sci U S A **105**(7): 2403-2408.
Wilkinson, B. L., P. E. Cramer, N. H. Varvel, E. Reed-Geaghan, Q. Jiang, A. Szabo, K. Herrup, B. T. Lamb and G. E. Landreth (2012). "Ibuprofen attenuates oxidative damage through NOX2 inhibition in Alzheimer's disease." Neurobiol Aging **33**(1): 197 e121-132.
Wosten, H. A. (2001). "Hydrophobins: multipurpose proteins." Annu Rev Microbiol **55**: 625-646.
Wuthrich, K. (1990). "Protein structure determination in solution by NMR spectroscopy." J Biol Chem **265**(36): 22059-22062.
Yamaguchi, K., H. Katou, M. Hoshino, K. Hasegawa, H. Naiki and Y. Goto (2004). "Core and heterogeneity of beta2-microglobulin amyloid fibrils as revealed by H/D exchange." J Mol Biol **338**(3): 559-571.
Yewdell, J. W., C. C. Norbury and J. R. Bennink (1999). "Mechanisms of exogenous antigen presentation by MHC class I molecules in vitro and in vivo: implications for generating CD8+ T cell responses to infectious agents, tumors, transplants, and vaccines." Adv Immunol **73**: 1-77.
Yin, J., X. Xia, Y. Shi, Y. Lu, C. Zhao, Z. Huang and N. Tian (2014). "Chinese familial transthyretin amyloidosis with vitreous involvement is associated with the transthyretin mutation Gly83Arg: a case report and literature review." Amyloid.
Zhang, R., X. Hu, H. Khant, S. J. Ludtke, W. Chiu, M. F. Schmid, C. Frieden and J. M. Lee (2009). "Interprotofilament interactions between Alzheimer's Abeta1-42 peptides in amyloid fibrils revealed by cryoEM." Proc Natl Acad Sci U S A **106**(12): 4653-4658.

Appendix

A) Amino acid sequences

CsgA [*E. coli*] – amino acid sequence (GenBank entry AAA23616)

1 MKLLKVAAIAAIVFSGSALAGVVPQYGGGGNHGGGGNNSGPNSELNIYQYGGGNSALALQ
61 TDARNSDLTITQHGGGNGADVGQGSDDSSIDLTQRGFGNSATLDQWNGKNSEMTVKQFGG
121 GNGAAVDQTASNSSVNVTQVGFGNNATAHQY

CsgB [*E. coli*]– amino acid sequence (GenBank entry CAA62281)

1 MKNKLLFMMLTILGAPGIAAAAGYDLANSEYNFAVNELSKSSFNQAAIIG QAGTNNSAQL
61 RQGGSKLLAVVAQEGSSNRAKIDQTGDYNLAYIDQAGSAN DASISQGAYG NTAMIIQKGS
121 GNKANITQYGTQKTAIVVQRQSQMAIRVTQR

B) All constructs carried a hexa-histidine tag, highlighted in Red.

$CsgA_{21\text{-}151}$ – amino acid sequence

20 MGVVPQYGGGGNHGGGGNNSGPNSELNIYQYGGGNSALALQTDARNSDLTITQHGGGNGA
80 DVGQGSDDSSIDLTQRGFGNSATLDQWNGKNSEMTVKQFGGGNGAAVDQTASNSSVNVTQ
140 VGFGNNATAHQYHHHHHH

$CsgB_{21\text{-}151}$ – amino acid sequence

20 MAGYDLANSEYNFAVNELSKSSFNQAAIIGQAGTNNSAQLRQGGSKLLAVVAQEGSSNRA
80 KIDQTGDYNLAYIDQAGSANDASISQGAYGNTAMIIQKGS GNKANITQYGTQKTAIVVQR
140 QSQMAIRVTQ RHHHHHH

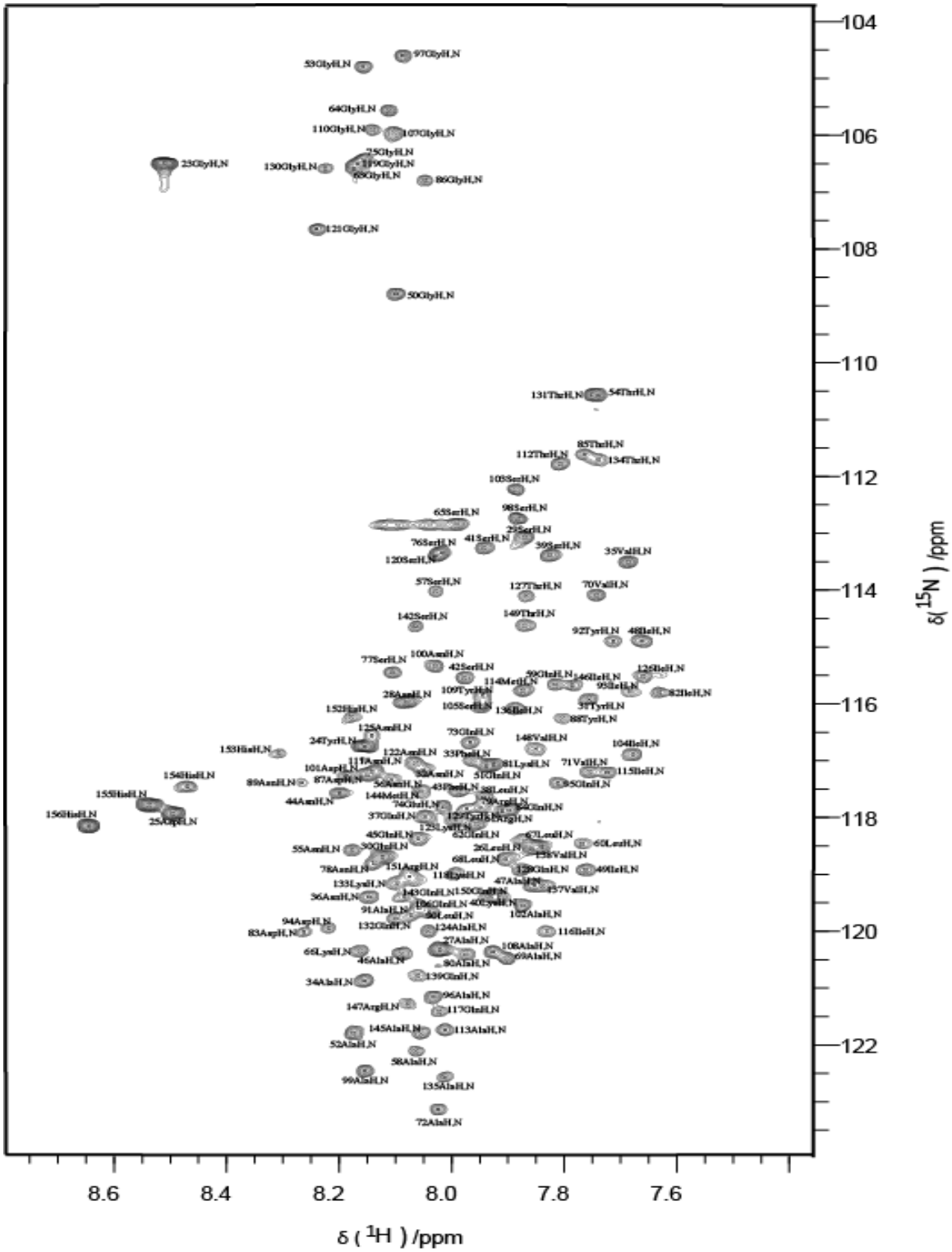
N-CGS-$CsgB_{21\text{-}151}$ – amino acid sequence

20 MCGSAGYDLANSEYNFAVNELSKSSFNQAAIIGQAGTNNSAQLRQGGSKLLAVVAQEGSSNRA
80 KIDQTGDYNLAYIDQAGSANDASISQGAYGNTAMIIQKGS GNKANITQYGTQKTAIVVQR
140 QSQMAIRVTQ RHHHHHH

List of Figures

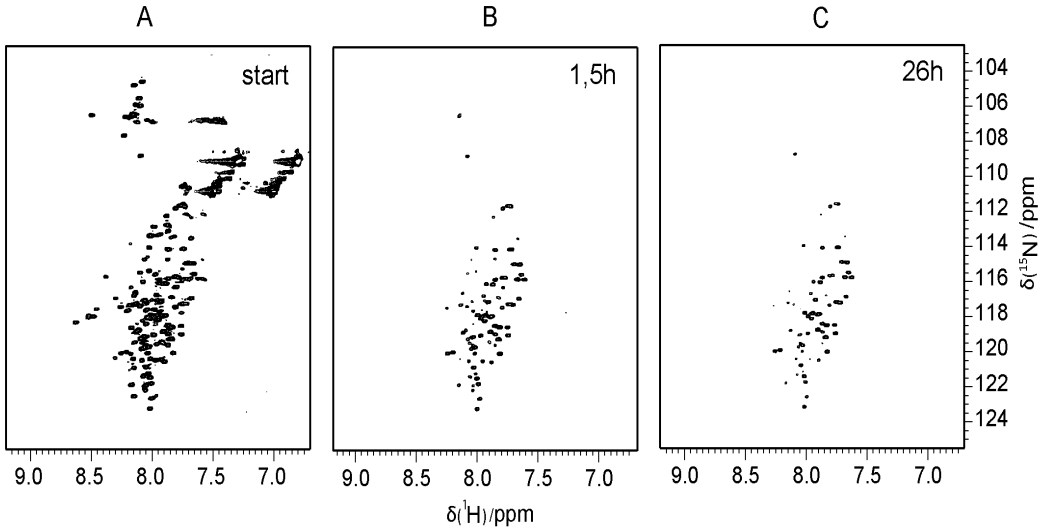
List of Tables

Supplementary Figures



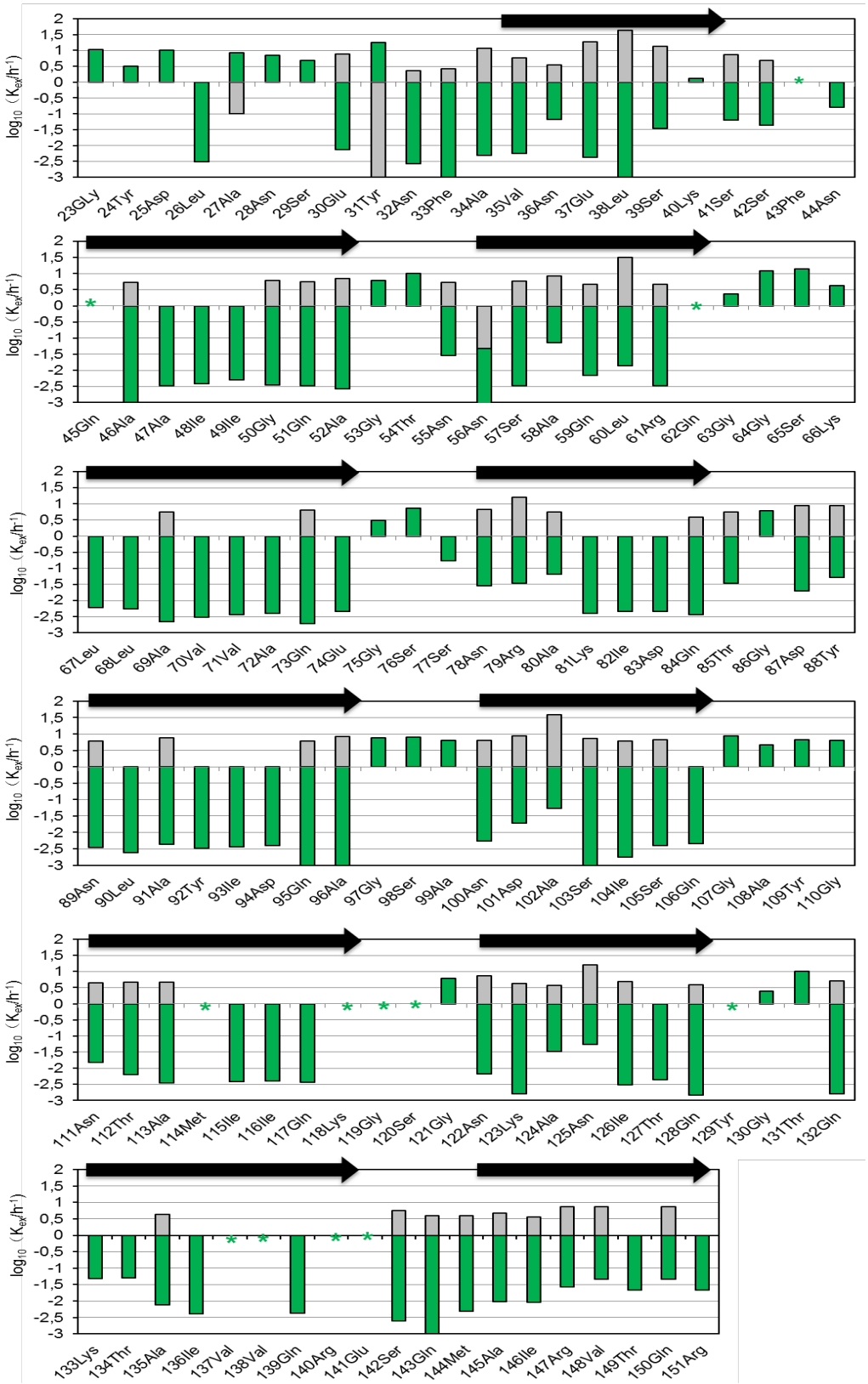
Supplementary figure 1: Backbone NH-assignments of CsgB

HSQC spectrum of uniformly ^{15}N labeled CsgB (20 mg wet pellet of fibrils) in d6-DMSO containing 0.1 % (v/v) d1-TFA. Sequence-specific chemical shift assignments are indicated next to the corresponding resonance.



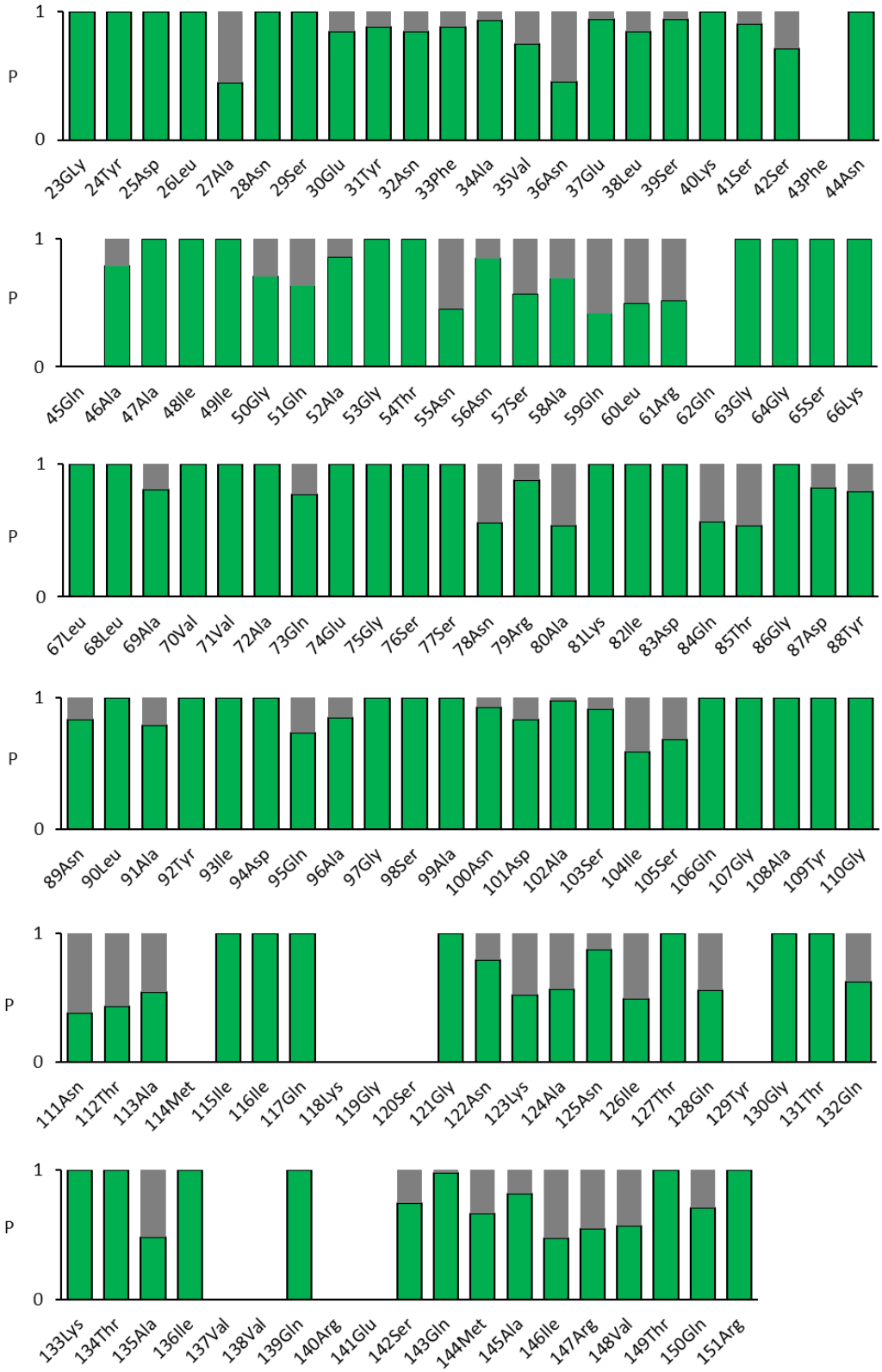
Supplementary figure 2: H/D exchange of CsgB fibrils.

HSQC spectra of CsgB fibrils (20 mg wet pellet of fibrils) dissolved in d6-DMSO containing 0.1 % (v/v) d1-TFA. **(A)** The spectra correspond to initial protonated state (start), **(B)** 1.5 hours exchange fibrils and **(C)** 26 hours exchange fibrils.



Supplementary figure 3: H/D exchange NMR data of CsgB in CsgB fibrils and secondary structure prediction

Green and grey bars indicated exchange rates (k_{ex} h^{-1}), gray bars indicated exchange rates of fast first phase in case of biphasic exchange behavior.residues marked with a green star were missing assigments.Black arrows indicated β-strands.



Supplementary figure 4: Relative populations in heterogenous fibers of CsgB in CsgB fibrils

Grey bars indicate the minor populations (fast exchange) and green indicate major populations (slow exchange) in case of biphasic behavior.

Curriculum Vitae

Personal information:

- Name: Puwei Yuan
- E-mail: puweiyuan@hotmail.com
- Born: 21.05.1982
- Place: Shannxi, China

Work experience and Education:

- **Helmholtz Centre for infection Research & TU Braunschweig** (07 2011-12 2014)
  PhD student

- **Institute for Molecular Medicine Finland, Helsinki University** (01 2010-02 2011)
  Researcher

- **ENN Group** (08 2008-12 2010)
  Researcher

- **Northwest University & Institute of Biophysics, Chinese Academy of Sciences** (09 2005-07 2008)
  Master Degree in Microbiology

- **Xi'an Technological University** (09 2001-07 2005)
  Bachelor degree in Polymer Chemistry

Publications:

1. Puwei Yuan, Mark Bartlam, Zhiyong Lou, Shoudeng Chen, Jie Zhou, Xiaojing He, Zongyang Lv, Ruowen Ge, Xuemei Li, Tao Deng, Ervin Fodor, Zihe Rao & Yingfang Liu. Crystal structure of an avian influenza polymerase PAN reveals an endonuclease active site. ***Nature***.**458**, 909-913. (2009)
2. Puwei Yuan, Jianxin Huang & Zhang Wang. Screening of High Cyclodextrin Glucanotransferase Productive Microbial Strains. ***Food and Fermentation Industries*** (Chinese). 33(**10**), 67-69. (2008)

Prize:

- First Prize, Beijing Science & Technology Award (2010)